

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 1871.78 Seconds

(without alignments) 10979.953 Million cell updates/sec

Title: US-09-674-379A-15

Perfect-score: 1269

Sequence: 1 cagtcagcagtcagtcgttga.....atgtgtcagtcagtcacatc 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estb1:*
2: em_estb1:*
3: em_estb1:*
4: em_estb1:*
5: em_estb1:*
6: em_estb1:*
7: em_estb1:*
8: em_estb1:*
9: em_estb1:*
10: em_estb1:*
11: em_estb1:*
12: em_estb1:*
13: em_estb1:*
14: em_estb1:*
15: em_estb1:*
16: em_estb1:*
17: em_estb1:*
18: em_estb1:*
19: em_estb1:*
20: em_estb1:*
21: em_estb1:*
22: em_estb1:*
23: em_estb1:*
24: em_estb1:*
25: em_estb1:*
26: em_estb1:*
27: em_estb1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	803.6	63.3	996	B0720516	B0720516 AGENCOURT
2	792.2	62.4	904	B0716569	B0716569 AGENCOURT
3	791	62.3	1014	BM921371	BM921371 AGENCOURT
4	771	60.8	954	BM545444	BM545444 AGENCOURT
5	666.8	52.5	891	BF339448	BF339448 AGENCOURT
6	665	52.4	822	BG109391	BG109391 602280448

7	639.2	50.4	796	13	BI453348	BI453348 603173067
8	636	50.1	644	13	BC938621	BC938621 CM27C11.X
9	592.8	46.7	717	13	BI251219	BI251219 602994427
10	576.4	45.4	933	14	BQ716088	BQ716088 AGENCOURT
11	573.6	45.2	627	12	BF090699	BF090699 QV2-NT004
12	566.2	44.6	882	13	BI825121	BI825121 603038662
13	565.2	44.5	870	14	BQ930927	BQ930927 AGENCOURT
14	564.4	44.5	912	13	BI766816	BI766816 603057118
15	549.6	43.3	880	13	BI764020	BI764020 603043221
16	544	42.9	835	14	BQ718885	BQ718885 AGENCOURT
17	540.2	42.6	627	13	BI183862	BI183862 UNL-P-FN-
18	537.8	42.4	745	12	BC431026	BC431026 602498636
19	510	40.2	531	10	AW957458	AW957458 EST369423
20	507.4	40.0	794	13	BI837271	BI837271 603090563
21	495.8	39.1	815	12	BI714360	BI714360 602336371
22	486	38.3	487	9	AL041285	AL041285 DKFZP34M
23	480	37.8	721	13	BI523975	BI523975 603052129
24	479	37.7	491	14	BM707581	BM707581 UT-E-C11-
25	477.8	37.7	494	14	BO695207	BO695207 1001350.H
26	459.4	36.2	1006	9	AL545549	AL545549 AL545549
27	446.2	35.2	762	13	BI834712	BI834712 603090304
28	442	34.8	690	12	BF104873	BF104873 601822595
29	415.4	32.7	576	10	AM142226	AM142226 EST292462
30	405.4	31.9	889	14	BO897144	BO897144 AGENCOURT
31	387.8	30.6	391	14	BM699401	BM699401 UT-E-DX1-
32	384.2	30.3	462	10	BE656312	BE656312 UT-M-BH0-
33	381.2	30.0	685	9	AL543659	AL543659 AL543659
34	377	29.7	629	13	BI489863	BI489863 603031610
35	375.4	29.6	377	14	N95751	N95751 zbf5e06.s1
36	374.2	29.5	1015	13	BM544363	BM544363 AGENCOURT
37	371.2	29.3	1850	11	BC018871	BC018871 Homo sapi
38	370.4	29.2	389	10	BE004365	BE004365 CM0-BN010
39	366.6	28.9	430	10	AW744572	AW744572 ut-32h12.y
40	359.2	28.3	906	14	BQ723059	BQ723059 AGENCOURT
41	358.2	28.2	443	9	AA801465	AA801465 EST190962
42	356	28.1	477	9	AA037185	AA037185 zc51c05.x
43	355.4	28.0	470	10	AW572176	AW572176 xt75e10.x
44	353	27.8	586	14	BM695168	BM695168 UT-E-CQ1-
45	349	27.5	371	12	BF090673	BF090673 QV2-NT004

ALIGNMENTS

RESULT 1
LOCUS B0720516 996 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT 8219503 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6186164 5', mRNA sequence.
ACCESSION B0720516
VERSION B0720516.1 GI:21859413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 996)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNML Unpublished (1999)
COMMENT Contact: Robert Strusberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13583 row: f column: 05
High quality sequence stop: 634.
Location/Qualifiers
1..996

FEATURES
SOURCE


```

Db      61 CCCGAGCCCTCCGAGAGACATGATGTTGTTAACCAAAATGGCGGATTTATGATC 120
Qy      133 CCCCCGAAACCCCTGTATGAGGGCCCTTCTGCAACCCCTACTGACCCCTACTCA 192
Db      121 CCCCCGAAACCCCTGTATGAGGGCCCTTCTGCAACCCCTACTGACCCCTACTCA 180
Qy      193 GGTCCGTCACGAGCTGCCCCACACCTCTGAGTCCCAACATCCCATCTCCAGG 252
Db      181 GGTCCGTCACGAGCTGCCCCACACCTCTGAGTCCCAACATCCCATCTCCAGG 240
Qy      253 CCTCTATATGCGCGTTTGATACGATGATGAAGAACCAATGTGTGATGTGAC 312
Db      241 CCTCTATATGCGCGTTTGATACGATGATGAAGAACCAATGTGTGATGTGAC 300
Qy      313 GAGTGTGAACAGATTTCCACAGTGCACACCCCAATCTGCATCAATACTGAAGC 372
Db      301 GAGTGTGAACAGATTTCCACAGTGCACACCCCAATCTGCATCAATACTGAAGC 360
Qy      373 GGTGACACCTCTCTCTGACCGACGAGATTTGGCTTCTGGAAGCCAGTCTTACAT 432
Db      361 GGTGACACCTCTCTCTGACCGACGAGATTTGGCTTCTGGAAGCCAGTCTTACAT 420
Qy      433 GATGAATGTGCTATGTTACTGACGACGCTCTGTGCAATGTTCTGATCTTATCT 492
Db      421 GATGAATGTGCTATGTTACTGACGACGCTCTGTGCAATGTTCTGATCTTATCT 480
Qy      493 TGTACATGACACCCGTTTACCTCAATGAGATGAGAAAGTCTTGCCAAATGTGAC 552
Db      481 TGTACATGACACCCGTTTACCTCAATGAGATGAGAAAGTCTTGCCAAATGTGAC 540
Qy      553 GAGTGTGACACGAGAACCCCTGCTGCAACCTGCTCAACACCTACGCTCTTTCATC 612
Db      541 GAGTGTGACACGAGAACCCCTGCTGCAACCTGCTCAACACCTACGCTCTTTCATC 600
Qy      613 TGGCGCTGTGACCCAGATGATGAACTTGAAGAGATGCGCTTCAATGCAATGAGAC 672
Db      601 TGGCGCTGTGACCCAGATGATGAACTTGAAGAGATGCGCTTCAATGCAATGAGAC 660
Qy      673 GAGTGTGACCTCTCTGATGCTCTGCAACATGATGTAAGACCGCCGACATAC 732
Db      661 GAGTGTGACCTCTCTGATGCTCTGCAACATGATGTAAGACCGCCGACATAC 720
Qy      733 TTTGCTCTCTGCTCTGACGCTACAT-CTGCTGTGATGACCAAC-CAAGCTGACAGA-- 788
Db      721 TTTGCTCTCTGCTCTGACGCTACAT-CTGCTGTGATGACCAACCGAAGCTGACAGA 780
Qy      789 CATCAAGCAATGTGACACAGGACCAACAGTGCAC-CTGCAAGCAAGCTGTACAT 847
Db      781 ATCAACGCAATGTGACACAGGACCAACAGTGCAC-CTGCAAGCAAGCTGTACAT 840
Qy      848 T---ACAAGGGGGGCTTCAATGATGACCCCATCGCGTGTGAGAG 891
Db      841 TTACAAGGGGGGCTTCAATGATGACCCCATCGCGTGTGAGAG 887

```

RESULT 3
 BM921371
 LOCUS 1014 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT.6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
 5', mRNA sequence.
 ACCESSION BM921371
 VERSION BM921371.1 GI:19371750
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1014)
 NIH-MGC <http://mgi.nci.nih.gov/>;
 AUTORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

```

FEATURES
source
1..1014
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752785"
/clone_1ib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."
BASE COUNT 235 a 300 c 246 g 233 t
ORIGIN
Query Match 62.3%; Score 791; DB 14; Length 1014;
Best Local Similarity 98.2%; Pred. No. 4.9e-205;
Matches 842; Conservative 0; Mismatches 10; Indels 5; Gaps 4;
Qy      1 CAGTGCACGAATGGCTTTGACCTGATCCGACGACGATGTTAGATTTGATGAA 60
Db      163 CAGTGCACGAATGGCTTTGACCTGATCCGACGACGATGTTAGATTTGATGAA 222
Qy      61 TCCCGAACCATCCCGAGGCTGCGGAGACATGATGTTAACAACCAATGGCGGG 120.
Db      223 TCCCGAACCATCCCGAGGCTGCGGAGACATGATGTTAACAACCAATGGCGGG 282
Qy      121 TATTTATGATTCCTCCGACCAACCTGTGTATGAGGGCCCTACTGCAACCCCTACTCG 180
Db      283 TATTTATGATTCCTCCGACCAACCTGTGTATGAGGGCCCTACTGCAACCCCTACTCG 342
Qy      181 ACCCCCTACTCAGGTCCGTAACCAAGACGCTGCCACACCTCTAGCTCCAACTATCCC 240
Db      343 ACCCCCTACTCAGGTCCGTAACCAAGACGCTGCCACACCTCTAGCTCCAACTATCCC 402
Qy      241 ACGATCTCCAGGCGCTTTATGATGCGGCTTGTGATACAGATGATGAAGCAACCAATGT 300
Db      403 ACGATCTCCAGGCGCTTTATGATGCGGCTTGTGATACAGATGATGAAGCAACCAATGT 462
Qy      301 GTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACCCACAGATGTGATC 360
Db      463 GTGATGTGACAGAGTGTGCAACAGATTTCCACAGTGCACACCCACCCACAGATGTGATC 522
Qy      361 AATACTGAAGGGGGGTACACCTCTCTGACCGACGAGATTTGGCTTTCTGGAAGCCAG 420
Db      523 AATACTGAAGGGGGGTACACCTCTCTGACCGACGAGATTTGGCTTTCTGGAAGCCAG 582
Qy      421 TGGTTAGACATGATGAATGTGCTATGTTACTGACGACGAGCTGTGCAATGTTTCT 480
Db      583 TGGTTAGACATGATGAATGTGCTATGTTACTGACGACGAGCTGTGCAATGTTTCT 642
Qy      481 GGATCCTATCTTGTACATGCAACCTGTGTTTACCTCAATGAGATGAGAGGTCTTGC 540
Db      643 GGATCCTATCTTGTACATGCAACCTGTGTTTACCTCAATGAGATGAGAGGTCTTGC 702
Qy      541 CAAGATGTAAAGAGTGTGCAACCGAGAACCCCTGCTGCAAACTGCTGAACACCTAC 600
Db      703 CAAGATGTAAAGAGTGTGCAACCGAGAACCCCTGCTGCAAACTGCTGAACACCTAC 762

```


Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LHAM9508 row: d column: 01
 High quality sequence stop: 675.
 Location/Qualifiers

FEATURES

SOURCE

1. 891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NCI-GAP_Brn64"
 /tissue_type="g10blastoma with EGFR amplification"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pCMV-SPO6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI-GAP Library."
 BASE COUNT 222 a 258 c 230 g 181 t
 ORIGIN

Query Match 52.5%; Score 666.8; DB 12; Length 891;
 Best Local Similarity 95.1%; Pred. No. 4e-171;
 Matches 742; Conservative 0; Mismatches 32; Indels 6; Gaps 5;

229 CCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGATGAA 288
 1 CCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGATGAA 59
 289 AGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGTGTCAACCCACC 348
 60 AGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGTGTCAACCCACC 119
 349 CAGATCTGCATCAATCACTGAAGGCGGTACACTGCTCTGACCCGAGGATATGGCTT 408
 120 CAGATCTGCATCAATCACTGAAGGCGGTACACTGCTCTGACCCGAGGATA-TGGCTT 178
 409 CTGGAAGGCGAGTGTGTAGACATGTGATGATGTGCTATGCTAGTCCAGACAGCTGTG 468
 179 CTGGAAGGCGAGTGTGTAGACATGTGATGATGTGCTATGCTAGTCCAGACAGCTGTG 238
 469 GCGAATGTCTGTGATCTTATCTTGTATGATGCAACCCGTGGTTTAACCTCAATGAGAT 528
 239 GCGAATGTCTGTGATCTTATCTTGTATGATGCAACCCGTGGTTTAACCTCAATGAGAT 298
 529 GGAAGGTCTGCGCAAGATGTGAGAGAGTGTGCAACCCGTGGTTTAACCTCAATGAGAT 588
 299 GGAAGGTCTGCGCAAGATGTGAGAGAGTGTGCAACCCGTGGTTTAACCTCAATGAGAT 358
 589 GTCAACACCTAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAAGAGAT 648
 359 GTCAACACCTAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAAGAGAT 418
 649 GGGCTTCATTCAGTATGATGTGAGAGAGTGTGAGTTCCTGTGCAACATGAG 708
 419 GGGCTTCATTCAGTATGATGTGAGAGAGTGTGAGTTCCTGTGCAACATGAG 478
 709 TGTGTGAACCGCCCGGACATCTTCTGCTGCGCCGACGAGTACATCTGCTGGAT 768
 479 TGTGTGAACCGCCCGGACATCTTCTGCTGCGCCGACGAGTACATCTGCTGGAT 538
 769 GACAACCGAGCTGTGCAAGACATCAACGATGTGAGACAGAACACAGTGTCAACCTG 828
 539 GACAACCGAGCTGTGCAAGACATCAACGATGTGAGACAGAACACAGTGTCAACCTG 598
 829 CAGCAGAGCTGTCAATTTTACAGGGGCTTCAAA-TGCATGACCCCATCCGCTGTGA 887
 599 CAGCAGAGCTGTCAATTTTACAGGGGCTTCAAA-TGCATGACCCCATCCGCTGTGA 658
 888 GGAGCTTATCTGAGAGATCAGTATACCGCT-GTATGTGCTGCTGAGAACCTTGGCT 946
 659 GGAGCTTATCTGAGAGATCAGTATACCGCTGAGTATGTGCTGCTGAGAACCT--TGGT 716
 947 GCAGAGACGACGCTTTACATCTTGTACCGGGACATGAGAGTGTGTGAGAGCTCCG 1006

Db 717 GCGAACAACAGCCCTTACCATCTGTCCGGACATGCGCCCTGGGGCAGAGCGCGCGC 776

RESULT 6
 BG109391
 LOCUS
 DEFINITION 822 bp mRNA linear EST 30-JAN-2001
 602280448F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4368227 5',
 mRNA sequence.
 BG109391
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC <http://imgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabds-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LHAM10021 row: 0 column: 12
 High quality sequence stop: 676.
 Location/Qualifiers

FEATURES

SOURCE

1. 822
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NIH_MGC_86"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: pCMV-SPO6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.533 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 204 a 227 c 210 g 181 t
 ORIGIN

Query Match 52.4%; Score 665; DB 12; Length 822;
 Best Local Similarity 97.8%; Pred. No. 1.2e-170;
 Matches 717; Conservative 0; Mismatches 10; Indels 6; Gaps 4;

493 TGTACATGCAACCTGCTTTTACCTCAATGAGATGAGAGTCTTGCCAAAGATGAA 552
 5 TGTACATGCAACCTGCTTTTACCTCAATGAGATGAGAGTCTTGCCAAAGATGAA 64
 553 GAGTGTGCCAGGAGAACCCCTGCGTCAACCTGCGTCAACACCTACGGCTTTTATC 612
 65 GAGTGTGCCAGGAGAACCCCTGCGTCAACCTGCGTCAACACCTACGGCTTTTATC 124
 613 TCCCGCTGTGACCAGATATGAGATGAGAGATGAGAGTCTTGCAAGATATGAG 672
 125 TCCCGCTGTGACCAGATATGAGATGAGAGATGAGAGTCTTGCAAGATATGAG 184
 673 GAGTGCAGCTTCTCTGAGTTCCTGTGCAACATGAGTGTGAAACGACCCGGCAGATAC 732
 185 GAGTGCAGCTTCTCTGAGTTCCTGTGCAACATGAGTGTGAAACGACCCGGCAGATAC 244
 733 TTCTGCTCTGCGCCGCTACATCTGCTGTGATGACAGAGTGTGCAAGATATGAG 792
 245 TTCTGCTCTGCGCCGCTACATCTCTGTGATGACAGAGTGTGCAAGATATGAG 304
 793 AACGATGTGAGACAGAACCAACAGTGTGAGAGTGTGAGAGTGTGATCAATTTACAA 852

Db 305 AACGAATGTGACGACAGAACACACAGCTGCACACCTGCAGACGAGCTCTACAAATTACAA 364
 QY 853 GGGGGCTTCAATGATCGACGCCATCCGCTGTGAGAGGCTATCGAGATCAGTAT 912
 Db 365 GGGGGCTTCAATGATCGACGCCATCCGCTGTGAGAGGCTATCGAGATCAGTAT 424
 QY 913 AACCGCTGTATGTCTCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 972
 Db 425 AACCGCTGTATGTCTCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 484
 QY 973 TACCGGACATGACGCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 1032
 Db 485 TACCGGACATGACGCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 544
 QY 1033 GCCAGCAGCCGCTACCGCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 1090
 Db 545 GCCAGCAGCCGCTACCGCTGTGAGAACCCCTGTGAGAGGCTATCGAGATCAGTAT 603
 QY 1091 GCAGAGATTTTACATGCGGCAAAAGGCGCCATCATGAGTGCACCTGTGATGACAGCC 1150
 Db 604 GCAGAGATTTTACATGCGGCAAAAGGCGCCATCATGAGTGCACCTGTGATGACAGCC 663
 QY 1151 CCATCAAA-GGGCGCGGGAATTCAGCTGAGTGGAAATGATCACTGTCAACACTGTC 1209
 Db 664 CCATCAAAAGGCGCGGGAATTCAGCTGAGTGGAAATGATCACTGTC--GCATGTC 721
 QY 1210 ATCACTGTCAGAG 1222
 Db 722 ATCAATTTCAAG 734

RESULT 7
 BI453348 796 bp mRNA linear EST 21-AUG-2001
 LOCUS 603173067F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5252369 5,
 DEFINITION mRNA sequence.
 ACCESSION BI453348
 VERSION BI453348.1 GI:15244004
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 796)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-femail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L14M1637 row: f column: 18
 High quality sequence stop: 796.

FEATURES
 source location/Qualifiers
 1..796
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5252369"
 /clone_1lb="NCI_CGAP_Mam5"
 /tissue_type="tumor, gross tissue"
 /dev_stage="7 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6; site:1: salI;
 site:2: NotI; cloned unidirectionally. Primer: oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Robin Humphreys,
 NIH"

BASE COUNT 195 a 208 c 197 g 196 t
 ORIGIN

Query Match 50.4%; Score 639.2; DB 13; Length 796;
 Best Local Similarity 89.2%; Pred. No. 1.3e-153;
 Matches 712; Conservativeness 0; Mismatches 83; Indels 3; Gaps 2;

QY 279 GATGATGAAGACCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTCACAGTGG 338
 Db 1 GATGATGAAGACCAACCAATGTGTGATGTGAGAGTGTGCAACAGATTCACAGTGG 60
 QY 339 CAACCCACCCAGATGTGATCAATATCAATGAGAGGCGGTACACCTGTCTGACCGAGG 398
 Db 61 CAACCCACCCAGATGTGATCAATATCAATGAGAGGCGGTACACCTGTCTGACCGAGT 120
 QY 399 ATATTGCTTCTGAGAGCCAGTCTTATGACATTTGATGAATGTGCTATGCTTACCA 458
 Db 121 GTACTGCTTCTGAGAGCCAGTCTTATGACATTTGATGAATGTGCTATGCTTACCA 180
 QY 459 GCAGCTGTGCGAATGTCTGATGATCTTATGATGATGATGATGATGATGATGATGAT 518
 Db 181 GCAGCTGTGCGAATGTCTGATGATCTTATGATGATGATGATGATGATGATGATGAT 240
 QY 519 CAATGAGATGAGAGTCTTGCAGAGATGAGAGTGTGCAACGAGAGTGTGCAACGAGAG 578
 Db 241 CAACGAGATGAGAGTCTTGCAGAGATGAGAGTGTGCAACGAGAGTGTGCAACGAGAG 300
 QY 579 GCAACCTGCTGCAACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 638
 Db 301 TCAAGCTGTGCAACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 639 TGAAGAGATGAGAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 Db 361 TGAAGAGATGAGAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 699 CCAACATGATGTGTGACCAAGCCGCGACATATCTGCTCTGCTTCCAGGCTACAT 758
 Db 421 TCAACATGATGTGTGACCAAGCCGCGACATATCTGCTCTGCTTCCAGGCTACAT 480
 QY 759 CCGTGTGATGACCAACGAGGCGGCAAGCATCAAGATGATGATGATGATGATGATGAT 818
 Db 481 CCGTGTGATGATGACCAAGGCGGCGGCAAGCATCAAGATGATGATGATGATGATGATGAT 540
 QY 819 GTGCAACCTGACAGAGAGTGTGATCAATTTACAAAGGCGGCTTCAATGATGATGAT 878
 Db 541 GTGCAACCTGACAGAGAGTGTGATCAATTTACAAAGGCGGCTTCAATGATGATGAT 600
 QY 879 CCGTGTGACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 937
 Db 601 CAGTGTGACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 938 ACCGTGTGACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 997
 Db 661 ACACCACTGTGACAGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 998 GAGCTGTGCTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1057
 Db 721 GAGCTGTGCTTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
 QY 1058 ATTACATTTTCCAGATCA 1075
 Db 779 ATTACATTTTCCAGATCA 796

RESULT 8
 LOCUS BG938621 644 bp mRNA linear EST 20-JUN-2002
 DEFINITION cn27c11.x1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
 clone NHTBC_cn27c11 random, mRNA sequence.
 ACCESSION BG938621
 VERSION BG938621.1 GI:14337993
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 644)
AUTHORS Jia, L.B., Young, M.F., Touchman, J.W., Bouffard, G.G.,
Beckstrom-Sternberg, S.M., Green, E.D., Powell, J.I., Yang, L.M., Robey
, P.G., Hotchkiss, R.N. and Francimano, C.A.
SCAP: The Skeletal Anatomy Project
JGAP: The Skeletal Anatomy Project
TITLE Unpublished (1997)
JOURNAL Contact: Libin Jia
COMMENT Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 27 row: c column: 11
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1. 644
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NHTBC-cn27c11"
/clone_1lb="Normal Human Trabecular Bone Cells"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/note="Organ: Hip; Vector: pBluescript; Site: 1; EORI:
Library constructed by Dr. Marian Young and Dr. Pamela
Gehron Robey (NIDR)." 149 t
146 a 195 c 154 g 149 t

BASE COUNT 146 a 195 c 154 g 149 t

ORIGIN

Query Match 50.1%; Score 636; DB 13; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.8e-163; Indels 0; Gaps 0;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

134 CCGGACAAACCCCTGTGTATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTCAG 193
|||||
9 CCGGACAAACCCCTGTGTATCGAGGGCCCTACTCGAACCCCTACTGACCCCTACTCAG 68
|||||
194 GTCCGTACCCAGAGCGCCCACTCTGACCTCCAACTATCCAGATCTCCAGCG 253
|||||
69 GTCCGTACCCAGAGCGCCCACTCTGACCTCCAACTATCCAGATCTCCAGCG 128
|||||
254 CTCTTATATGCGCTTTGGATACAGATGATGAAAGCAACCAATGTGTGATGTGAGC 188
|||||
129 CTCTTATATGCGCTTTGGATACAGATGATGAAAGCAACCAATGTGTGATGTGAGC 373
|||||
314 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 373
|||||
189 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 248
|||||
374 GGTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 433
|||||
249 GGTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 308
|||||
434 ATGAATGCTGCTATGCTTACTGCGACAGCTCTGTGCGAATGTTCTGATCTTATCTT 493
|||||
309 ATGAATGCTGCTATGCTTACTGCGACAGCTCTGTGCGAATGTTCTGATCTTATCTT 368
|||||
494 ATGAATGCTGCTATGCTTACTGCGACAGCTCTGTGCGAATGTTCTGATCTTATCTT 553
|||||
369 GTATATGCAACCCCTGTTTACCTCAATGAGATGGAAGGCTTCCAGATGTGAAG 428
|||||
554 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 613
|||||
429 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 488
|||||
614 GCGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 673

DB 489 GCGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 548
|||||
OY 674 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 733
|||||
DB 549 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 608
|||||
OY 734 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769
|||||
DB 609 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
|||||

RESULT 9
B1251219 717 bp mRNA linear EST 17-JUL-2001
LOCUS B1251219
DEFINITION 602994427F1 NCL-CCAP_Mam5 Mus musculus cDNA clone IMAGE:5150191 5',
B1251219 mRNA sequence.
ACCESSION B1251219 GI:14800410
VERSION B1251219.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 717)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@biml.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
http://image.llnl.gov
Plate: LLNL1371 row: e column: 08
High quality sequence stop: 713.
Location/Qualifiers
1. 717
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5150191"
/clone_1lb="NCL-CCAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; SalI;
Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lohar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 172 a 186 c 182 g 177 t

ORIGIN

Query Match 46.7%; Score 592.8; DB 13; Length 717;
Best Local Similarity 89.8%; Pred. No. 5.9e-151; Indels 0; Gaps 0;
Matches 636; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

251 GCGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 310
|||||
1 GCGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 60
|||||
DB 311 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 370
|||||
OY 61 AGTGGCAACAGATTCACCAAGTGAACACCCACCCAGATCTGATCAATCTGAAGCG 120
|||||
DB 371 GCGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 430
|||||
OY 121 GAGGTTACACCTGCTCTGACACCGAGATATGCTTCTGGAAGGCCAGTCTTAACATG 180
|||||

```

OY 431 TTGATGAATGTCCTATGTTACTGTCACAGCTCTGTGAGATGTCCTGATCTAT 490
    |||||||
Db 181 TTGATGAATGTCCTATGTTACTGTCACAGCTCTGTGAGATGTCCTGATCTAT 490
OY 491 CTGTGATGACAAACCTGTTTACCTCAATGAGAGTGAAGGTCTGGCAAGATGTA 240
    |||||||
Db 241 CCTGTACAAACCAACCTGTTTACCTCAATGAGAGTGAAGGTCTGGCAAGATGTA 240
OY 551 ACAGTGTGCCACAGAAACCCCTGCTGCAACACCTGCTGCAACACCTGCTGCTTCA 610
    |||||||
Db 301 ACAGTGTGCCACAGAAACCCCTGCTGCAACACCTGCTGCTGCTGCTGCTTCA 610
OY 611 TCTGCGCTGTGACCCAGATATGACTGAGAGATGAGTGTGCTGCTGCTGCTTCA 360
    |||||||
Db 361 TCTGCGCTGTGACCCAGATATGACTGAGAGATGAGTGTGCTGCTGCTGCTTCA 360
OY 671 ACAGTGTGCCACAGAAACCCCTGCTGCAACACCTGCTGCTGCTGCTGCTTCA 420
    |||||||
Db 421 ACAGTGTGCCACAGAAACCCCTGCTGCAACACCTGCTGCTGCTGCTGCTTCA 420
OY 731 ACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 480
    |||||||
Db 481 ACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 480
OY 791 TCACGAATGTGAGCAGACAGAACACAGTGTGCTGCTGCTGCTGCTGCTTCA 540
    |||||||
Db 541 TCACGAATGTGAGCAGACAGAACACAGTGTGCTGCTGCTGCTGCTGCTTCA 540
OY 851 AAGGGGGCTTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 600
    |||||||
Db 601 AAGGGGGCTTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 600
OY 911 ATAAACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 958
    |||||||
Db 958 ATAAACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 958
OY 661 AAAACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 708
    |||||||

```

RESULT 10
 LOCUS B0716088
 DEFINITION AGENCOURT_8099803 Lupski_1_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6190063 5', mRNA sequence.
 B0716088
 B0716088.1 GI:21854985
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
 NIH-MGC http://mgi.nci.nih.gov/.
 CONTACT: Robert Strausberg, Ph.D.
 E-mail: c9apbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM13588 row: e column: 08
 High quality sequence stop: 501.
 Location/Qualifiers
 1..933
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone IMAGE:6190063
 /clone_1db="Lupski_sympathetic_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"

```

/lab_host="DH10B"
/Note="Vector: PCW-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
5'-raccaccacccgccc-3' and
5'-gactagctttgagtcgagagcggccgcccct(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT 225 a 284 c 245 g 179 t
ORIGIN

```

```

Query Match 45.4%; Score 576.4; DB 14; Length 933;
Best Local Similarity 94.6%; Pred. No. 2e-146;
Matches 619; Conservative 0; Mismatches 31; Indels 4; Gaps 2;

```

```

OY 527 ATGAAGGCTTGGCAAGATGTGAAGAGTGTGCAACGAAACCCCTGCTGCAACCT 586
    |||||||
Db 2 ATGAAGGCTTGGCAAGATGTGAAGAGTGTGCAACGAAACCCCTGCTGCAACCT 586
OY 587 GCCTGCAACACCTGAGGCTCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 61
    |||||||
Db 62 GCCTGCAACACCTGAGGCTCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 61
OY 647 ATGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 121
    |||||||
Db 122 ATGCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 121
OY 707 AGTGTGTGAACAGCCCGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 181
    |||||||
Db 182 AGTGTGTGAACAGCCCGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 181
OY 767 ATGACAAACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGAT 241
    |||||||
Db 242 ATGACAAACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGAT 241
OY 827 TGCAGCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 301
    |||||||
Db 302 TGCAGCAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCA 301
OY 887 AGGAGCTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
    |||||||
Db 362 AGGAGCTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
OY 947 GCAGAGACAGCCCTTTACCATCTGTACCGGACATGAGCTGTGCTGCTGCTGCT 421
    |||||||
Db 422 GCAGAGACAGCCCTTTACCATCTGTACCGGACATGAGCTGTGCTGCTGCTGCT 421
OY 1007 TTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTTACCTGCGCTATTACATT 1066
    |||||||
Db 1066 TTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTTACCTGCGCTATTACATT 1066
OY 482 TTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTTACCTGCGCTATTACATT 541
    |||||||
Db 541 TTCCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTTACCTGCGCTATTACATT 541
OY 1067 TCCAGATCAATCTGGGAATGA-66GCGAGAGAAATTTTTCATGCGGCAAA--CGGGCCCC 1122
    |||||||
Db 1122 TCCAGATCAATCTGGGAATGA-66GCGAGAGAAATTTTTCATGCGGCAAA--CGGGCCCC 1122
OY 542 TCCAGATCAATCTGGGAATGA-66GCGAGAGAAATTTTTCATGCGGCAAA--CGGGCCCC 601
    |||||||
Db 601 TCCAGATCAATCTGGGAATGA-66GCGAGAGAAATTTTTCATGCGGCAAA--CGGGCCCC 601
OY 1123 ATCAGTGTCCACCTGAGTATGAGACAGCCCATCAAGAGGCGCGGGAATCCAG 1176
    |||||||
Db 1176 ATCAGTGTCCACCTGAGTATGAGACAGCCCATCAAGAGGCGCGGGAATCCAG 1176
OY 602 ATCAGTGTCCACCTGAGTATGAGACAGCCCATCAAGAGGCGCGGCGGCGG 655
    |||||||

```

```

FEATURES
Source
    .RESULT 11
    LOCUS BF090699/c
    DEFINITION OY2-NT0048-120900-360-a11 NT0048 Homo sapiens cDNA, mRNA sequence.
    ACCESSION BF090699
    VERSION BF090699.1 GI:10896409
    KEYWORDS EST.
    SOURCE human.
    ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1 (bases 1 to 627)	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
	Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordi, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	20202663	
	Contact: Simpson A.J.G.				
	Laboratory of Cancer Genetics				
	Ludwig Institute for Cancer Research				
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil				
	Tel: +55-11-2704922				
	Fax: +55-11-2707001				
	Email: asimpson@ludwig.org.br				
	This sequence was derived from the FAPESP/LICR Human Cancer Genome				
	http://www.ludwig.org.br/scripts/gethtml2.pl?tl=8t2-QV2-NT0048-1202				
	900-360-alle63-2000-09-12&ct4=1)				
	Seq primer: puc 18 forward				
	High quality sequence start: 10				
	High quality sequence stop: 606.				
FEATURES					
SOURCE					
	1. 627				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_id="NT0048"				
	/dev_stage="Adult"				
	/note="Organ: nervous_tumor; Vector: puc18; Site 1: Smal; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Patents Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
BASE COUNT	148 a 161 c 170 g 148 t				
ORIGIN					
Query Match	45.2%; Score 573.6; DB 12; Length 627;				
Best Local Similarity	98.2%; Freq. No. 9.7e-146;				
Matches 591; Conservative	0; Mismatches 9; Indels 2; Gaps 1;				
Oy	369 AGCGGAGTACACCTGCTCTGACACGAGATATGCTTCGGAAGCCAGTCTTGA 428				
Db	613 AAGCGGGACACCTGCTCTGACACGAGATATGCTTCGGAAGCCAGTCTTGA 554				
Oy	429 CATTGATGATGCGCTATGCTTACTGCGACGACGCTGTGCGAATGTTCTCGATCTA 488				
Db	553 CATTGATGATGCTATGCTTACTGCGACGACGCTGTGCGAATGTTCTCGATCTA 494				
Oy	489 TTCTTGTCATGACCAACCTGGTTTTTACCCTCAATAGAGATGGAAGGCTTGCCAAAGATG 548				
Db	493 TTCTTGACATGACCAACCTGGTTTTTACCCTCAATAGAGATGGAAGGCTTGCCAAAGATG 434				
Oy	549 GAAGAGATGTCACACCGAAGACCCCTGGTGCACAACTGCGTCAACACCTTACGCTCTTT 608				
Db	433 GAAGAGATGTCACACCGAAGACCCCTGGTGCACAACTGCGTCAACACCTTACGCTCTTT 374				
Oy	609 CATCTGCCGCTGTGACCCAGATATGACACTTGAGAAATGGCGTTCAATGCAAGATAT 668				
Db	373 CATCTGCCGCTGTGACCCAGATATGACACTTGAGAAATGGCGTTCAATGCAAGATAT 314				
Oy	669 GGAGAGATGACAGCTCTCTGATGCTCTGCAACATGATGTTGCAACAGCCGGGAC 728				
Db	313 GGAGAGATGACAGCTCTCTGATGCTCTGCAACATGATGTTGCAACAGCCGGGAC 254				
Oy	729 ATACTTGTCTCTGCTCCCTCCAGGCTACATCTCTGATGACAAACGAGCTGCCAAGA 788				

Db		253	ATATTCTTGCCTCCGCTCAGGCTACATCTCTGGATTACAACCAACTGCCAGA	194
OY		789	CATAAAGCATGTGTAGACAGAGAACCAACAGCTGCAACTTGAGAGACAGTCCTACAAATT	848
Db		193		
OY		193	CATTAACGAATGTATAGACAGAGAACCAACAGCTGCAACTTGAGAGACAGTCCTACAAATT	134
Db		849	ACAAAGGGGGCTTCAAATGATCGACCCCATCGCTGTGAGAGAGCCTTATCTGAGAGATCAG	908
OY		909	TGATTAACCGCTGTATGTGTCTCTGCTGAGAACCTTGCTGCGAGAGACCAGCCCTTTACAT	968
Db		73		
OY		73	TGATTAACCGCTGTATGTGTCTCTGCTGAGAACCC--TGTCGAGAGACAGCCCTTTACAT	16
Db		133	ACAAAGGGGGCTTCAAATGCGTTAGCCCATCCGCTGTGAGAGAGCCTTATCTGAGAGATCAG	74
OY		969	CT 970	
Db		15	CT 14	
RESULT 12				
B1825121				
LOCUS		882 bp	mRNA	linear EST 04-OCT-2001
DEFINITION		60303866ZP1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5'		
ACCESSION		B1825121		
VERSION		B1825121.1		
KEYWORDS		GI:15936671		
SOURCE		EST.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE		NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: cgabs-remail.nih.gov		
		Tissue Procurement: Life Technologies, Inc.		
		cDNA Library Preparation: Life Technologies, Inc.		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LML at:		
		http://image.llnl.gov		
		Plate: LHAM11447 row: 1 column: 15		
		High quality sequence start: 14		
		High quality sequence stop: 844.		
FEATURES		Location/Qualifiers		
source		1..882		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone_image="5179550"		
		/clone_id="NIH_MGC_115"		
		/lab_host="DH10B"		
		/note="Organ: pooled brain, lung, testis; Vector:		
		PCMV-SPORE; Site.1: NotI; site.2: EcoRV (destroyed); RNA		
		source anonymous pool of 6 male brains, age range 23-27; 1		
		male lung, age 27; and 1 male testis, age 69. Library is		
		oligo-dt primed and directionally cloned (EcoRV site is		
		destroyed upon cloning). Average insert size 1.8 kb,		
		insert size range 1-3 kb. Library is normalized and		
		enriched for full-length clones and was constructed by C.		
		Gruher (Invitrogen). Research Genetics tracking code		
		021. Note: this is a NIH_MGC Library "		
BASE COUNT		191 a 277 c 216 g 197 t		1 others
ORIGIN				
Query Match		44.6%; Score 566.2; DB 13; Length 882;		
Best Local Similarity		97.9%; Pred. NO. 1.2e-143;		
Matches		615; Conservative 0; Mismatches 9; Indels 4; Gaps 4;		
OY		1	CAGTGCACGAATGGCTTGACCTGGATGCCAGTAGACAGACGTGTTAGATTGATGATA	60


```

Db      254 CAGTGCAGCAATGGCTTTACCTGGATGCCCACTGACGACAGTGTAGATATGATGAA 313
QY      61 TCGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 120
Db      314 TGCCGACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 373
QY      121 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 180
Db      374 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 433
QY      181 ACCCCTACTGAGTCCGACGACGAGCGCCCACTGAGTCCCAACTATGCC 240
Db      434 ACCCCTACTGAGTCCGACGACGAGCGCCCACTGAGTCCCAACTATGCC 493
QY      241 ACAGTCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 300
Db      494 ACAGTCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 553
QY      301 TGGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 360
Db      554 GTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 613
QY      361 AATACTGAAGGCGGTACACCTGCTCTGACGACGAGATATGCTTGTGGAAGGCCAG 420
Db      614 AATACTGAAGGCGGTACACCTGCTCTGACGACGAGATATGCTTGTGGAAGGCCAG 673
QY      421 TGGCTAGACATGATGATGCTGCTATGCTTACTGCGAGAGCTGTGGAAGTCTCT 480
Db      674 TGGCTAGACATGATGATGCTGCTATGCTTACTGCGAGAGCTGTGGAAGTCTCT 733
QY      481 GGATCTATCTTGTACATGCAACCCCT-GGTTTACCTCAATGAGATGAG-AGGTCCT 538
Db      734 GGATCTATCTTGTACATGCAACCCCTGGGTTTACCTCAATGAGATGAGAGTCTT 793
QY      539 GCCAAGATGTGAAG-AGTGTGCCACCGAGAACCCCTGCGTGAACCTGCGTCAACAC 597
Db      794 GCCAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGCGTGAACCTGCGTCAACAC 853
QY      598 TACGGCTCTTCACTGCGCGCTGTGACC 625
Db      854 TA-GGGCTTTCTCATCTGCGCGCTGTGACC 880

RESULT 13
BO930927 870 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_8955548 NCL_CGAP_Mam2 Mus musculus cDNA clone
DEFINITION IMAGE:6439732 5', mRNA sequence.
ACCESSION BO930927.1 GI:22345958
VERSION EST.
KEYWORDS Mus musculus
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabps-femail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13960 row: d column: 05
High quality sequence stop: 651.
Location/Qualifiers
FEATURES
source 1..870
/organism="Mus musculus"

```

```

Query Match 44.5%; Score 565.2; DB 14; Length 870;
Best Local Similarity 87.9%; Pred. No. 2.3e-143;
Matches 649; Conservative 0; Mismatches 85; Indels 4; Gaps 3;

Db      1 CAGTGCAGCAATGGCTTTGACCTGATCGCCAGTCAAGTGTATGATATGATGAA 60
QY      123 CAGTGCAGCAATGGCTTTGACCTGATCGCCAGTCAAGTGTATGATATGATGAA 182
Db      61 TCGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 120
QY      183 TCGCGAACCATCCCGAGGCGCTGCGGAGAGACATGATGTGTAAACAAATGGCGGG 242
Db      121 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 180
QY      243 TATTATGATTCCTCCCGGACAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTCG 302
Db      181 ACCCCTACTGAGTCCGACGACGAGCTGCCCCACCTGCTCAAGTCAACCTATGCC 240
QY      303 ACATCTACTGAGGCGCATACCCAGAGCGGCGCCACAGTACAGCTTCACCACTACCC 362
Db      241 ACAGTCTCCAGGCGCTTATATGCGGCTTTGATACCATGATGATGAAGCAACCAATGT 300
QY      363 ACATCTACTGAGGCGCATACCCAGAGCGGCGCCACAGTACAGCTTCACCACTACCC 422
Db      423 GTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 482
QY      301 TGGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 360
Db      423 GTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 482
QY      361 AATACTGAAGGCGGTACACCTGCTCTGACGACGAGATATGCTTGTGGAAGGCCAG 420
Db      483 AACACTGAAGAGAGTGTGCAACAGATTCGCCACAGTGAACCCCAAGATGCGATC 542
QY      421 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db      543 TGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY      481 GGATCTATCTTGTACATGCAACCCCTGGTTTACCTCAATGAGATGAGAGTCTTGC 540
Db      603 GGATCTATCTTGTACATGCAACCCCTGGTTTACCTCAATGAGATGAGAGTCTTGC 662
QY      541 CAAGATGTGAAGAGTGTGCAACAGATTCGCCACAGTGAACCCCTGCGTGAACCTGAC 600
Db      663 CAAGATGTGAAGAGTGTGCAACAGATTCGCCACAGTGAACCCCTGCGTGAACCTGAC 722
QY      601 GGCTCTTCAATCTGCGCTGTGACCAAGATATG-ACTTGAAGAGATGCGGCTCA-TT 658
Db      723 GGCTCTTCAATCTGCGCTGTGACCAAGATATGAACTTGAAGAGATGCGGCTCA-TT 782
QY      659 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
Db      783 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
QY      717 CCAAGCCGCGACATATCT 734
Db      843 ACAAGCCGCGCTCATATCTT 860

RESULT 14
BIT68516

```


Query Match	44.5%	Score 564.4	DB 13	Length 912
Best Local Similarity	99.8%	Pred. No. 3.8e+143		
Matches	565	Conservative	0	Mismatches 1
			Indels	Gaps 0
QY	704	ATGAGTGTGTGAACACGAGCCCGGACATCTCTGCTCTGCTGCTTCAGAGCTACATCCGCG	763	
Db	1	ATGAGTGTGTGAACACGAGCCCGGACATCTCTGCTCTGCTGCTTCAGAGCTACATCCGCG	60	
QY	764	TGATGTACAACCGAAGCTGCCAGACATCAACGAATGTGAGCACAGAACACAGTGCA	823	
Db	61	TGATGTACAACCGAAGCTGCCAGACATCAACGAATGTGAGCACAGAACACAGTGCA	120	
QY	824	ACCGAGCAGAGAGTGTGCTACAATTTACAGGGGGCTTCAATTCGATCGACCCCATCCGCT	883	
Db	121	ACCGAGCAGAGAGTGTGCTACAATTTACAGGGGGCTTCAATTCGATCGACCCCATCCGCT	180	
QY	884	GTGAGGAGCCCTTATCTGAGATCAGTGATTAACCCCTGTATGTCTCTGAGAACCTGT	943	
Db	181	GTGAGGAGCCCTTATCTGAGATCAGTGATTAACCCCTGTATGTCTCTGAGAACCTGT	240	
QY	944	GCTCAGAGACACGAGCCCTTTACCATCTTTGTACCCGGAGATGAGCGTGTGACGAGCT	1003	
Db	241	GCTCAGAGACACGAGCCCTTTACCATCTTTGTACCCGGAGATGAGCGTGTGACGAGCT	300	
QY	1004	CCGTTCCCGGTGAGATCTTCAAAATGCAAGGACGAGACCCGGTATACCTTGGGCGCTATTACA	106	
Db	301	CCGTTCCCGGTGAGATCTTCCAAATGCAAGGACGAGACCCGGTATACCTTGGGCGCTATTACA	360	

```

1. .880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5183693"
/clone_1ib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPORT6; Site_1: NOTI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Institrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT      192 a      278 c      215 g      195 t
ORIGIN

Query Match      43.3%; Score 549.6; DB 13; Length 880;
Best Local Similarity 98.3%; Pred. No. 4.1e-139;
Matches 619; Conservative 0; Mismatches 4; Indels 7; Gaps 6;

1 CAGTCGACGANTGGCTTGACCTGGATGCCAGTCGACGACGATGTTGATATGATGAA 60
|||||
252 CAGTCGACGANTGGCTTGACCTGGATGCCAGTCGACGACGATGTTGATATTCATGAA 311

```

```
OY 61 TCGGAAACCATCCCGAGGCTGCCGAGAGACATGATGTGTATACCAAAATGCGGG 120
   |||||||
Db 312 TCGGAAACCATCCCGAGGCTGCCGAGAGACATGATGTGTATACCAAAATGCGGG 371
   |||||||
OY 121 TATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 180
   |||||||
Db 372 TATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 431
   |||||||
OY 181 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATGCC 240
   |||||||
Db 432 ACCCCCTACTGAGTCCGTACCCAGAGCTGCCACACTCTCAGCTCCAAACTATGCC 491
   |||||||
OY 241 ACGATCTCCAGGCTCTTATATGCCCTTGTGATACCAATGATGAAAGCAACCAATGT 300
   |||||||
Db 492 ACGATCTCCAGGCTCTTATATGCCCTTGTGATACCAATGATGAAAGCAACCAATGT 551
   |||||||
OY 301 GTGATGTGACAGAGTGTGCAACAGATTGCCACAGTGCACCCACCCAGATCTGATC 360
   |||||||
Db 552 GTGATGTGACAGAGTGTGCAACAGATTGCCACAGTGCACCCACCCAGATCTGATC 611
   |||||||
OY 361 AATACTGAAGCGGCTACACCTGCTCTGACACGAGGATATGGCTTCTGGAAGGCCAG 420
   |||||||
Db 612 AATACTGAAGCGGCTACACCTGCTCTGACACGAGGATATGGCTTCTGGAAGGCCAG 671
   |||||||
OY 421 TGCTTAGACATTTGATGATGTCTGCTATGCTTACTGCTGCGCAGCA -GCTCTGAGGATGTTCC 479
   |||||||
Db 672 TGCTTAGACATTTGATGATGTCTGCTATGCTTACTGCTGCGCAGCA -GCTCTGAGGATGTTCC 731
   |||||||
OY 480 TGATCTTATTTCTTTGA -CATGCAACCCCTGGTTTACCTCAATGAGGATGAA -GGTC 536
   |||||||
Db 732 TGATCTTATTTCTTTGA -CATGCAACCCCTGGTTTACCTCAATGAGGATGAA -GGTC 791
   |||||||
OY 537 TTGCCAAGA -TGAGACGAGTGTGCGACCGAGAACCCCTGGCTGCAAACTGGCTCAACA 595
   |||||||
Db 792 TTGCCAAGA -TGAGACGAGTGTGCGACCGAGAACCCCTGGCTGCAAA -CTGGCTCAACA 849
   |||||||
OY 596 CCTACGGCTCTTTCATCTGCGCGCTGTGACC 625
   |||||||
Db 850 CCTACGGCTCTTTCATCTGCGCGCTGTGACC 879
   |||||||
```

Search completed: July 3, 2003, 17:36:08
Job time : 1873.78 secs

QY	483	CACGATCTCCAGGCGCTCTTATATGCGGCTTTGGATTACCAGATGGATGAAGCAACCAATG	542
Db	764	CACGATCTCCAGGCGCTCTTATATGCGGCTTTGGATTACCAGATGGATGAAGCAACCAATG	823
QY	543	TGTGGATGTGGACAGATGTGCAAGATATCCACAGTGCAGAACCCCAACCCAGATCTGCAT	602
Db	824	TGTGGATGTGGACAGATGTGCAAGATATCCACAGTGCAGAACCCCAACCCAGATCTGCAT	883
QY	603	CAATACTGAAGGCGGGGTACACCTGCTCCTGCACGACGAGATATTGGCTTCTGGAAGGCCA	662
Db	884	CAATACTGAAGGCGGGGTACACCTGCTCCTGCACGACGAGATATTGGCTTCTGGAAGGCCA	943
QY	663	GTGCTTGGACATTTATGAATGTCCTATNGTTACTTGCACAGAGCTGTGGGATATTCC	722
Db	944	GTGCTTGGACATTTATGAATGTCCTATNGTTACTTGCACAGAGCTGTGGGATATTCC	1003
QY	723	TGATTCCTATCTTGTATCATGCAACCCGTGTATTACCTCATAGATAGGATGAAGTCTTG	782
Db	1004	TGATTCCTATCTTGTATCATGCAACCCGTGTATTACCTCATAGATAGGATGAAGTCTTG	1063
QY	783	CCAAGATGTGAACAGATGTGCCACCCGAGAACCCCTGGGTGCAAACTTGCCTCAACCTTA	842
Db	1064	CCAAGATGTGAACAGATGTGCCACCCGAGAACCCCTGGGTGCAAACTTGCCTCAACCTTA	1123
QY	843	CGGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTAGGAAGATGGCGTTCAATG	902
Db	1124	CGGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTAGGAAGATGGCGTTCAATG	1183
QY	903	CAGTGATATGGACAGATGACAGCTTCTCTGAATTCCTCTGCCACATGATGTGTGAACA	962
Db	1184	CAGTGATATGGACAGATGACAGCTTCTCTGAATTCCTCTGCCACATGATGTGTGAACA	1243
QY	963	GCCGGGACATATCTGCTGCTGCTGCTCCGACAGGCTACATCTGCTGGATGCAACCCGAAG	1022
Db	1244	GCCGGGACATATCTGCTGCTGCTGCTCCGACAGGCTACATCTGCTGGATGCAACCCGAAG	1303
QY	1023	CTGGCAAGACATCAACGATGTGACACAGAGAACACACACGTGCAACCTGCACGACAGAGTG	1082
Db	1304	CTGGCAAGACATCAACGATGTGACACAGAGAACACACACGTGCAACCTGCACGACAGAGTG	1363
QY	1083	CTACATATTTACAAGGGGCTTCAAAATGATGACAGCCCATCCGCTGTAGAGAGCCTTATCT	1142
Db	1364	CTACATATTTACAAGGGGCTTCAAAATGATGACAGCCCATCCGCTGTAGAGAGCCTTATCT	1423
QY	1143	GAGGATCAGTATTAACCGCTGTATGTGCTCTGTGAGAAACCTCGGTGCAAGACCAAGCC	1202
Db	1424	GAGGATCAGTATTAACCGCTGTATGTGCTCTGTGAGAAACCTCGGTGCAAGACCAAGCC	1483
QY	1203	CTTTACCATCTTGTACGGGGACATGAGAGTGTGTCAAGAGCTTCCTGCTCCGCTGACAT	1262
Db	1484	CTTTACCATCTTGTGTACGGGGACATGAGAGTGTGTCAAGAGCTTCCTGCTCCGCTGACAT	1543
QY	1263	CTTTCCAATGCAAGCCACAGACCCGCTACCCCTGGGGCCATTACATATTTCCAGATCAATC	1322
Db	1544	CTTTCCAATGCAAGCCACAGACCCGCTACCCCTGGGGCCATTACATATTTCCAGATCAATC	1603
QY	1323	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCGCACCTGGT	1382
Db	1604	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGCGCACCTGGT	1663
QY	1383	GATGACACGCCCCCATCAAAAGGGGCCCGGGAATCCAGCTGGACATTTGGAAATGATCACCTGT	1442
Db	1664	GATGACACGCCCCCATCAAAAGGGGCCCGGGAATCCAGCTGGACATTTGGAAATGATCACCTGT	1723
QY	1443	CAACACTGTCTCAACTTTCAGAGGACAGCTCCGCTATCCATGCACTGGGAGATATGTGTGCA	1502
Db	1724	CAACACTGTCTCAACTTTCAGAGGACAGCTCCGCTATCCATGCACTGGGAGATATGTGTGCA	1783
QY	1503	GTACCCATTTGAGAGCTTGCGGCTGTGAGCTCTCCAGACGTGCTCTCATTTGGCACCAAGGGA	1562
Db	1784	GTACCCATTTGAGAGCTTGCGGCTGTGAGCTCTCCAGACGTGCTCTCATTTGGCACCAAGGGA	1843

QY	1563	CAGAGAAAGAGAGAAATTAACAGAGAAATGAGAGACACAGACAGAGCTTAGCAATTTCTG	1622
Dp	1844	CAGAGAAAGAGAGAAATTAACAGAGAAATGAGAGACACAGACAGAGCTTAGCAATTTCTG	1903
QY	1623	CTGAACGTTTCCCAGAGAGTCACGCCGACTTCCTGACTCTCACCTGACTATTTCAGA	1682
Dp	1904	CTGAACGTTTCCCAGAGAGTCACGCCGACTTCCTGACTCTCACCTGACTATTTCAGA	1963
QY	1683	CTGTCAACCTGCAGAGACTTGGCACCCCGAGTTCCTATGATACAGTTATCAAAAAGTATT	1742
Dp	1964	CTGTCAACCTGCAGAGACTTGGCACCCCGAGTTCCTATGATACAGTTATCAAAAAGTATT	2023
QY	1743	ATCATTCCTCCCCGATAGAAAGATTGTGTGGAAATTTTCAAGGCCCTCAGTTTATTCCA	1802
Dp	2024	ATCATTCCTCCCCGATAGAAAGATTGTGTGGAAATTTTCAAGGCCCTCAGTTTATTCCA	2083
QY	1803	CTATTTTCAAAAGAAATAGATTAGTTTGGGGGGGTGTGAGTCTATGTTCAAAAGACTGTG	1862
Dp	2084	CTATTTTCAAAAGAAATAGATTAGTTTGGGGGGGTGTGAGTCTATGTTCAAAAGACTGTG	2143
QY	1863	AACAGCTTGCTGTACATTCTTACCTTCTCACCTCTTCTCTACCTGTGTTACTGCTTTG	1922
Dp	2144	AACAGCTTGCTGTACATTCTTACCTTCTCACCTCTCTCTCACCTGTGTTACTGCTTTG	2203
QY	1923	CAAGAGCCCGGGAGCTGGGGGGAAACCTGGGAGTACTGTAGTTGGCTTTTGGCTACACA	1982
Dp	2204	CAAGAGCCCGGGAGCTGGGGGGAAACCTGGGAGTACTGTAGTTGGCTTTTGGCTACACA	2263
QY	1983	GAGAAAGCTATGTATAAACAAACACAGCAGATCCAAAGGTTTATGAGAAATGTTTCAA	2042
Dp	2264	GAGAAAGCTATGTATAAACAAACACAGCAGATCCAAAGGTTTATGAGAAATGTTTCAA	2323
QY	2043	AACCATGCTGGTATATTTTCAACCATATAAGAGTTTCAGTTGTCTTAATTTGTATPAC	2102
Dp	2324	AACCATGCTGGTATATTTTCAACCATATAAGAGTTTCAGTTGTCTTAATTTGTATPAC	2383
QY	2103	GGTTTAATTCGTCCTGTTCATTTTGGATTTTAAAAAATATGTGGTGAATTTCCCTTC	2162
Dp	2384	GGTTTAATTCGTCCTGTTCATTTTGGATTTTAAAAAATATGTGTGAATTTCCCTTC	2443
QY	2163	GAAAGGCTTCAGACACATGCTATGTTCTGTCTCCCAAAACCCAGTCTCCCTCCATATTT	2222
Dp	2444	GAAAGGCTTCAGACACATGCTATGTTCTGTCTCCCAAAACCCAGTCTCCCTCCATATTT	2503
QY	2223	AGCCCACTGTTTTTTTGGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAATT	2282
Dp	2504	AGCCCACTGTTTTTTTGGAGACCCCTTAATCTTGCTTTCTTTTGAATTTTACCCAATT	2563
QY	2283	GGATTGGAATGCAGAGGCTCCAAACGATTAAATTTTGAAGAGA 2328	
Dp	2564	GGATTGGAATGCAGAGGCTCCAAACGATTAAATTTTGAAGAGA 2609	

Search completed: July 3, 2003, 17:53:38
Job time : 382.284 secs

Db 1784 GTACCCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTGCCCTTCATTGGCACCACCAAGGGA 1843

OY	1143	GAGATCATGTAATTAACCGCTGTATGTGTCTCTGTAAGAACCCCTGGCTGCAGAGACCAAGC	1202
Db	1424	GAGGATCAGTGTAAACCCGCTGTATGTGTCTCTGTAAGAACCCCTGGCTGCAGAGACCAAGC	1483
OY	1203	CTTACACATCTGTACCGGACATGTGACAGCGTGTCAAGGACGCTCCGTTCCCGCTGACAT	1262
Db	1484	CTTACACATCTGTACCGGACATGTGACAGCGTGTCAAGGACGCTCCGTTCCCGCTGACAT	1543
OY	1263	CTTCCAAATGCAAGCCACGACCCGCTAACCTTGGGGCCTATTAATTTTCCAGATCAAAATC	1322
Db	1544	CTTCCAAATGCAAGCCACGACCCGCTAACCTTGGGGCCTATTAATTTTCCAGATCAAAATC	1603
OY	1323	TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGCCCAATCACTGGCACCCCTGT	1382
Db	1604	TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGCCCAATCACTGGCACCCCTGT	1663
OY	1383	GATGACAGGCCCCATCAAAAGGGCCCCGGGAATTCAGCTGACCTTGGAAATGATCACTGT	1442
Db	1664	GATGACAGGCCCCATCAAAAGGGCCCCGGGAATTCAGCTGACCTTGGAAATGATCACTGT	1723
OY	1443	CAACACTGTATCAATCTTCAGAGGACCTCCGTATCCGATCCGATGCGGAGATATATGTGCGCA	1502
Db	1724	CAACACTGTATCAATCTTCAGAGGACCTCCGTATCCGATCCGATGCGGAGATATATGTGCGCA	1783
OY	1503	GTACCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTGCTCTCATTTGGCACCAAGGA	1562
Db	1784	GTACCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTGCTCTCATTTGGCACCAAGGA	1843
OY	1563	CAGGAGAAAGAGGAAATTAACAGAGAAATGAGAGCGACACAGAGGTAATTCGATTCGTA	1622
Db	1844	CAGGAGAAAGAGGAAATTAACAGAGAAATGAGAGCGACACAGAGGTAATTCGATTCGTA	1903
OY	1623	CTGAAGCTTCCCGGAGAGTCAAGCCCGACTCTCTGACTCTCACCTGTACTATTTCAGA	1682
Db	1904	CTGAAGCTTCCCGGAGAGTCAAGCCCGACTCTCTGACTCTCACCTGTACTATTTCAGA	1963
OY	1683	CCTGTACCTCTGAGACCTTGCACACCCCACTTCTATGATACAGTATCAAAAGATAT	1742
Db	1964	CCTGTACCTCTGAGACCTTGCACACCCCACTTCTATGATACAGTATCAAAAGATAT	2023
OY	1743	ATCATATGCTCCCCGATAGAAAGATGTGTGGTGAATTTTCAAGGCCCTTCAGTTATTTCCA	1802
Db	2024	ATCATATGCTCCCCGATAGAAAGATGTGTGGTGAATTTTCAAGGCCCTTCAGTTATTTCCA	2083
OY	1803	CTATTTTCAAGAAATAGATTAGGTTTGGCGGGCTGTAGATGCTATGTTCAAGACGTG	1862
Db	2084	CTATTTTCAAGAAATAGATTAGGTTTGGCGGGCTGTAGATGCTATGTTCAAGACGTG	2143
OY	1863	AACAGCTGCTGACCTTCTCACCTCTCCACTCTCTCTCACCTGTGTACTGCTTG	1922
Db	2144	AACAGCTGCTGACCTTCTCACCTCTCCACTCTCTCTCTCACCTGTGTACTGCTTG	2203
OY	1923	CAAAAGCCGGAGCTGGCGGGGAAACCTGGGAGTAGTATGTTTGGCTTTTGGGTACACA	1982
Db	2204	CAAAAGCCGGAGCTGGCGGGGAAACCTGGGAGTAGTATGTTTGGCTTTTGGGTACACA	2263
OY	1983	GAAAGGCTATGTAAACAACAACAGACAGAGATCGAAGGGTTTTTAAAGATGTCTTCAA	2042
Db	2264	GAAAGGCTATGTAAACAACAACAGACAGAGATCGAAGGGTTTTTAAAGATGTCTTCAA	2323
OY	2043	AACCATGCTGTATTTTCAACCATTAAGAAAGTTCACTGTCTCTTAATTTGTATTAAC	2102
Db	2324	AACCATGCTGTATTTTCAACCATTAAGAAAGTTCACTGTCTCTTAATTTGTATTAAC	2383
OY	2103	GGTTTAATTCGTCTTGTTCATTTTGAGTATTTTAAAAAATATGTGCGTAATTCCTTC	2162
Db	2384	GGTTTAATTCGTCTTGTTCATTTTGAGTATTTTAAAAAATATGTGCGTAATTCCTTC	2443
OY	2163	GAAAGGCTTACAGACATGCTATGTCTGTCTTCCCAACCAAGTCTCCATTCATTTT	2222
Db	2444	GAAAGGCTTACAGACATGCTATGTCTGTCTTCCCAACCAAGTCTCCATTCATTTT	2503

OY	2223	AGCCGACGTCTTTCCTTTGAGGACCCCTTAACCTGCTTCTTGGAATTTTACCACATT	2282
Db	2504	AGCCGAGCGTTTCTTTGGAGAGCCCCCTAACTCTGCTTTTGAAATTTTACCACATT	2503
OY	2283	GGATTGGAATCAGAGGCTCTCCAACACTGATTAAATATTGGAAGA	2328
Db	2564	GCATTGGAATCAGAGGCTCTCCAACCTGATTAAATATTGGAAGA	2609
 RESULT 15 US-10-180-557-407 Sequence 407, Application US/10180557 Publication No. US20030022301A1 GENERAL INFORMATION: APPLICANT: Baker, Kevin P. APPLICANT: Chen, Jian APPLICANT: Desnoyers, Luc APPLICANT: Goddard, Audrey APPLICANT: Godowski, Paul J. APPLICANT: Gurney, Austin L. APPLICANT: Pau, James APPLICANT: Smith, Victoria APPLICANT: Watanabe, Colin K. APPLICANT: Wood, William I. APPLICANT: Zhang, Zemin TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC FILE REFERENCE: P3430R1C147 CURRENT APPLICATION NUMBER: US/10/180, 557 PRIORITY FILING DATE: 2002-06-25 Prior Application removed - See File Wrapper or Palm NUMBER OF SEQ ID NOS: 612 SEQ ID NO 407 LENGTH: 2609 TYPE: DNA ORGANISM: Homo Sapien US-10-180-557-407			
 Query Match 99.8%; Score 2322.8; DB 9; Length 2609; Best Local Similarity 99.9%; Pred. No. 0; Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0			
OY	3	CCCGGGGCTCTCCCGGTGTCCTCTCCACAGACTCGTGGGCCCTCTGGAATAAACACCC	62
Db	284	CCCGGGGCTCTCCCGGTGTCCTCTCCACAGACTCGTGGGCCCTCTGGAATAAACACCC	343
OY	63	GCGAGGCCCGAGGGCCCAAGAGAGGCCGAGAGTCCCGAGCTCTCCGGGGGTCGCCCGG	122
Db	344	GCGAGGCCCGAGGGCCCAAGAGAGGCCGAGAGTCCCGAGCTCTCCGGGGGTCGCCCGG	403
OY	123	CAGAGCTTCTTCGCGCTTCGATCTCTCCCTCGGCGCGTCTTGGAATGCCAGGAATAA	182
Db	404	CAGAGCTTCTTCGCGCTTCGATCTCTCCCTCGGCGCGTCTTGGAATGCCAGGAATAA	463
OY	183	AAGGATACACTACGTTACCATTCCTGCTCTGTTCACAAGCCCTGGGAATGCACAGGC	242
Db	464	AAGGATACACTACGTTACCATTCCTGCTCTGTTCACAAGCCCTGGGAATGCACAGGC	533
OY	243	ACAGTGCAGAAAGCGTTTGACCTGGATGCGCACGTCAAGACAGTGTTAGATATTGATGA	302
Db	524	ACAGTGCAGAAAGCGTTTGACCTGGATGCGCACGTCAAGACAGTGTTAGATATTGATGA	583
OY	303	ATGCCGAACATCCCGAGGCGTCGCGAGAGACATGATGTGTTAACCAAAATGGCGG	362
Db	584	ATGCCGAACATCCCGAGGCGTCGCGAGAGACATGATGTGTTAACCAAAATGGCGG	643
OY	363	GTAATTATGCAATTCCTCCGGAACAAACCTGTGTATCGAGGGCCCTACTCTGAAACCCCTACTC	422
Db	644	GTAATTATGCAATTCCTCCGGAACAAACCTGTGTATCGAGGGCCCTACTCTGAAACCCCTACTC	703
OY	423	GACCCCTACTCAAGGTCCGTAACCCAGAGGTGCCCCACCACTCTCAGCTCCAAACTATCC	482
Db	704	GACCCCTACTCAAGGTCCGTAACCCAGAGGTGCCCCACCACTCTCAGCTCCAAACTATCC	766

QY	1803	TTATTTTCAAGAAAATAGATAGCTTGGGGGCTGAGTCTATCTTCAACACGTG	1862
Db	2084	CTATTTTCAAGAAAATAGATAGCTTGGGGGCTGAGTCTATCTTCAACACGTG	2143
QY	1863	AACAGCTTGTCTACCTCTTCACCTCTCCACTCTCTCACTGTGTACTGCTTG	1922
Db	2144	AACAGCTTGTCTACCTCTTCACCTCTCCACTCTCTCTCACTGTGTACTGCTTG	2203
QY	1923	CAAGACCCGGAGCTGGCGGGAAACCTGGGATCTAGTTGCTTTTGCTTACACA	1982
Db	2204	CAAGACCCGGAGCTGGCGGGAAACCTGGGATCTAGTTGCTTTTGCTTACACA	2263
QY	1983	GAGAAGCTATGTATTAACAAACACACAGAGATCATCAAGGTTTTTGAAGATGTCTCAA	2042
Db	2264	GAGAAGCTATGTATTAACAAACACACAGAGATCATCAAGGTTTTTGAAGATGTCTCAA	2323
QY	2043	AACCATGCTGCTATTTTTCACCAATTAAGAGTTTCACTGTCCTTAATTTGTATAC	2102
Db	2324	AACCATGCTGCTATTTTTCACCAATTAAGAGTTTCACTGTCCTTAATTTGTATAC	2383
QY	2103	GGTTAAATTCGTCTGTCTCATTTTGATATTTTAAAAAATATGTGTGAATTCCTTC	2162
Db	2384	GGTTAAATTCGTCTGTCTCATTTTGATATTTTAAAAAATATGTGTGAATTCCTTC	2443
QY	2163	GAAAGGCTTCAGACATGCTATGTCTGTCTCCAAACCCAGCTCCTCTCCATTTT	2222
Db	2444	GAAAGGCTTCAGACATGCTATGTCTGTCTCCAAACCCAGCTCCTCTCCATTTT	2503
QY	2223	AGCCAGTGTTCCTTTGAGAGACCCCTTAATCTCTGCTTCTTTGAATTTTACCACAT	2282
Db	2504	AGCCAGTGTTCCTTTGAGAGACCCCTTAATCTCTGCTTCTTTGAATTTTACCACAT	2563
QY	2283	GGATTTGGAATGCAGAGTCTCCAAACGATTAATATTGGAAGCA 2328	
Db	2564	GGATTTGGAATGCAGAGTCTCCAAACGATTAATATTGGAAGCA 2609	

RESULT 14
US-10-180-552-407
; Sequence 407, Application US/10180552
; Publication No. US20030022300A1

```

1  APPLICANT: Baker, Kevin P.
2  APPLICANT: Chen, Jian
3  APPLICANT: Desnoyers, Luc
4  APPLICANT: Goddard, Audrey
5  APPLICANT: Godowski, Paul J.
6  APPLICANT: Gurney, Austin L.
7  APPLICANT: Pan, James
8  APPLICANT: Smith, Victoria
9  APPLICANT: Watanabe, Collin K.
10 APPLICANT: Wood, William I.
11 APPLICANT: Zhang, Zemin
12
13 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
14
15 FILE REFERENCE: P3430RIC153
16
17 CURRENT APPLICATION NUMBER: US/10/0180.552
18
19 PRIORITY FILING DATE: 2002-06-25
20
21 Prior Application removed - See File Wrapper or Palm
22
23 NUMBER OF SEQ ID NOS: 612
24
25 SEQ ID NO 407
26
27 LENGTH: 2609
28
29 TYPE: DNA
30
31 ORGANISM: Homo Sapien
32
33 US-10-180-552-407

```

	Query Match	99.8%	Score 2322.8;	DB 9;	Length 2609;
	Best Local Similarity	99.9%	Pred. No. 0;		
	Matches 2324; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;
07	3	CCCCGGCTCCCCGTCCTCTTCACAGACTGCTCGCCCCCTTGGAAATTAACACC	62		

Db	284	CCCGGCGGTCTCCCCGTTGCTCTCCACAGATGCTGCGGCCCTCTGTGAATTAACACACC	343
QY	63	GGAGGCCCCGAGGGGCCACAGAGAGGGCGAGGTGCCCCAGTCTCTCGGGGGTCCCCGCC	122
Db	344	GGGAGCCCCCGAGGGCCACAGAGAGGGCGAGGTGCCCCAGTCTCTCGGGGGTCCCCGCC	403
QY	123	CGACCTTCTCTCGCCCTTGCAATCTCTCTCTCGCGGGTCTTGGAATGCGAGAAATAA	182
Db	404	CGACCTTCTCTCGCCCTTGCAATCTCTCTCTCGCGGGTCTTGGAATGCGAGAAATAA	463
QY	183	AAGATACCTCACTGTACCATTTGAGTCTCTCTGTCTTCCAAAGCCCTGGAAATGCACAGC	242
Db	464	AAGATACTCACTGTACCATTTGAGTCTCTCTGTCTTCCAAAGCCCTGGAAATGCACAGC	523
QY	243	ACACTGCACGAATGGCTTTGACCGGAAATGCGAAATGCGACAGATGTTAGATATGATGA	302
Db	524	ACAGTGCACGAATGGCTTTGACCTGGATGCGCAATGCGCAATGAGATGTTAGATATGATGA	583
QY	303	ATGCGGAAACCATCCCGAGGGCTCCGAGAGAGACATGATGTGTAAACCAATATGGCGG	362
Db	584	ATGCGGAAACCATCCCGAGGGCTCCGAGAGAGACATGATGTGTAAACCAATATGGCGG	643
QY	363	GTATTTATGCAATCCCCGGGAAACCCGTGTATACGAGGGCCCTACCTCGAATCCCTCTCTC	422
Db	644	GTATTTATGCAATCCCCGGGAAACCCGTGTATACGAGGGCCCTACCTCGAATCCCTCTCTC	703
QY	423	GACCCCTCACTCAGGTGCGTACCCAGAGAGTGCACCACTCTCAGTCTCAACTATATCC	482
Db	704	GACCCCTCACTCAGGTGCGTACCCAGAGAGTGCACCACTCTCAGTCTCAACTATATCC	763
QY	483	CACGATCTCCAGGCGCTCTTATATGCGCTTTGGATACCAATGATGATGAACCAACAATG	542
Db	764	CACGATCTCCAGGCGCTCTTATATGCGCTTTGGATACCAATGATGATGAACCAACAATG	823
QY	543	TGTGAGATGAGACGAGTGTGCACAGATATCCACAGTGCACCAACCACCCAGATCTCAT	602
Db	824	TGTGAGATGAGACGAGTGTGCACAGATATCCACAGTGCACCAACCACCCAGATCTCAT	883
QY	603	CAATACGTGAAGGGGGGTACACCTCTCTCTGCACGCGAGGATATTTGGCTTGTGAAGGCCA	662
Db	884	CAATACGTGAAGGGGGGTACACCTCTCTCTGCACGCGAGGATATTTGGCTTGTGAAGGCCA	943
QY	663	GTGCTTGACATGATGAATGGAATGCGCTATGATGATTCAGACAGATCTGTGGATATATCC	722
Db	944	GTGCTTGACATGATGAATGGAATGCGCTATGATGATTCAGACAGATCTGTGGATATATCC	1003
QY	723	TGATATCCATTTCTGTACATGCAACCCGTGTATACCTCAATAGATGAGATGAGTCTTG	782
Db	1004	TGATATCCATTTCTGTACATGCAACCCGTGTATACCTCAATAGATGAGATGAGTCTTG	1063
QY	783	CCAAAGATGGAACGAGTGTGCCACGAGAACCCCTGGTGCAAACTGTGCTCAACACCTTA	842
Db	1064	CCAAAGATGGAACGAGTGTGCCACGAGAACCCCTGGTGCAAACTGTGCTGCTCAACACCTTA	1123
QY	843	CGGCTCTTCAATCTGCGCGCTGTGACCCGAGATATGAATTTAGGAAGAATGGCGTTCAATG	902
Db	1124	CGGCTCTCTCAATCTGCGCGCTGTGACCCGAGATATGAATTTAGGAAGAATGGCGTTCAATG	1183
QY	903	CAGTGAATGAGCAGTGCACGCTTCTGTGATGCTCTGCGCACATGCAAGTGTGAACA	962
Db	1184	CAGTGAATGAGCAGTGCACGCTTCTGTGATGCTCTGCGCCACATGATGATGTGAACA	1243
QY	963	GGCGGCAACATCTTGCTGCTCTGCGCCGCGAGGTACATCTCTGGTGGATGCAACCGAAG	1022
Db	1244	GGCGGCAACATCTTGCTGCTCTGCGCCGCGAGGTACATCTCTGGTGGATGCAACCGAAG	1303
QY	1023	CTGCGCAAGACATCAAGATGTGAGCAACAGAAACACACGTGCAACCTGCGACGACAGCTG	1082
Db	1304	CTGCGCAAGACATCAAGATGTGAGCAACAGAAACACACGTGCAACCTGCGACGACAGCTG	1363
QY	1083	CTACAATTTACAAAGGGGCTTCAAAATGATGACACCCCATCCGCTGTGAGAGGCTTATCT	1142
Db	1364	CTACAATTTACAAAGGGGCTTCAAAATGATGACACCCCATCCGCTGTGAGAGGCTTATCT	1423

APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C66
CURRENT FILING DATE: 2002-06-20
PRIORITY APPLICATION removed - see file wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-913-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3 CCGGGGCTCTCCCGCTGCTCTCCAGACGCTGCGGCGCCCTGGAAATMAAACACCC 62
284 CCGGGGCTCTCCCGCTGCTCTCCAGACGCTGCGGCGCCCTGGAAATMAAACACCC 343
63 GCGAGCCCGAGAGGCCAGAGAGAGCGAGCTGCGGCGCCCTGGAAATMAAACACCC 122
344 GCGAGCCCGAGAGGCCAGAGAGAGCGAGCTGCGGCGCCCTGGAAATMAAACACCC 403
123 CGAGCTTTTCTTCCGCTTTCGATCTCTCTGCGGCTCTGGACATGCGAGAAATAA 182
404 CGAGCTTTTCTTCCGCTTTCGATCTCTCTGCGGCTCTGGACATGCGAGAAATAA 463
183 AAGGATCTACGTTACATCTGCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 242
464 AAGGATCTACGTTACATCTGCTGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 523
243 ACAGTGACAGAAATGCGCTTTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 302
524 ACAGTGACAGAAATGCGCTTTCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583
303 ATGCCGACCAATCCCGAGGCTTCGAGAGAGATGATGCTGTTAAACCAAAATGCGCG 362
584 ATGCCGACCAATCCCGAGGCTTCGAGAGAGATGATGCTGTTAAACCAAAATGCGCG 643
363 GTATTATGCAATCCCGAGGCTTCGAGAGAGATGATGCTGTTAAACCAAAATGCGCG 422
644 GTATTATGCAATCCCGAGGCTTCGAGAGAGATGATGCTGTTAAACCAAAATGCGCG 703
423 GACCCCTCTACTAGTCCGATACCGAGAGCTGCGGCGCCCTGGAAATMAAACACCC 482
704 GACCCCTCTACTAGTCCGATACCGAGAGCTGCGGCGCCCTGGAAATMAAACACCC 763
483 CACGATCTCCAGGCTCTTATATGCGCTTTCGATACAGATGATGAAAGCAACCAATG 542
764 CACGATCTCCAGGCTCTTATATGCGCTTTCGATACAGATGATGAAAGCAACCAATG 823
543 TGTGGATGAGAGAGTGTGCAACAGATGCGGCGCCCTGGAAATMAAACACCC 602
824 TGTGGATGAGAGAGTGTGCAACAGATGCGGCGCCCTGGAAATMAAACACCC 883
603 CAATCTGAAGGCGGTACACCTGCTCTGCTGCGAGAGATGATGCTTTCGAAAGGCA 662
884 CAATCTGAAGGCGGTACACCTGCTCTGCTGCGAGAGATGATGCTTTCGAAAGGCA 943
663 GTGCTTAGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722

944 GTGCTTAGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
723 TGTGATCTTATCTTGTATCATGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
1004 TGTGATCTTATCTTGTATCATGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063
783 CCAAGATGTGAACAGAGTGTGCGACCGAGAACCCCTGCTGCAAACTGCTGCAACCTA 842
1064 CCAAGATGTGAACAGAGTGTGCGACCGAGAACCCCTGCTGCAAACTGCTGCAACCTA 1123
843 CGGCTCTTATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
1124 CGGCTCTTATCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183
903 CAGTGAATGAGAGAGTGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 962
1184 CAGTGAATGAGAGAGTGTGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243
963 GCCCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
1244 GCCCGGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303
1023 CTGCGAAGACATCAACGATGTGACACAGGACACACGCTGCAACCTGCTGCAACCTG 1082
1304 CTGCGAAGACATCAACGATGTGACACAGGACACACGCTGCAACCTGCTGCAACCTG 1363
1083 CTCAATTTTCAAGGCGGCTTCAAAATGATGACCCCTGCTGCTGCTGCTGCTGCTGCT 1142
1364 CTCAATTTTCAAGGCGGCTTCAAAATGATGACCCCTGCTGCTGCTGCTGCTGCTGCT 1423
1143 GAGATCAGTGAATACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
1424 GAGATCAGTGAATACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
1203 CTTTACCATCTTGTACCGGAGATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
1484 CTTTACCATCTTGTACCGGAGATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543
1263 CTTTCAATTTCAAGGCGGCTTCAAAATGATGACCCCTGCTGCTGCTGCTGCTGCTGCT 1322
1544 CTTTCAATTTCAAGGCGGCTTCAAAATGATGACCCCTGCTGCTGCTGCTGCTGCTGCT 1603
1323 TGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCGCCATCACTGCAACCTG 1382
1604 TGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCGCCATCACTGCAACCTG 1663
1383 GATGACACCGCCATCAAGGCGGCGGGAATCAGTGGACTGGAATGATGATCTG 1442
1664 GATGACACCGCCATCAAGGCGGCGGGAATCAGTGGACTGGAATGATGATCTG 1723
1443 CAACACTGTATCAACTTCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
1724 CAACACTGTATCAACTTCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1783
1503 GTACCATCTGAGCTCTGCGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
1784 GTACCATCTGAGCTCTGCGGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1843
1563 CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
1844 CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1903
1623 CTGAAGCTTCCCGAGAGAGTGTGCAACAGATGCGGCGCCCTGGAAATMAAACACCC 1682
1904 CTGAAGCTTCCCGAGAGAGTGTGCAACAGATGCGGCGCCCTGGAAATMAAACACCC 1963
1683 CCGTGAACCTGTGAGAGCTTGCACACCCAGTCTCTATGATGATGATGATGATGATGAT 1742
1964 CCGTGAACCTGTGAGAGCTTGCACACCCAGTCTCTATGATGATGATGATGATGATGAT 2023
1743 ATCATTTCTCCCTGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1802
2024 ATCATTTCTCCCTGATGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2083

```

|||||
Db 524 ACAGTGCAGAAATGGCTTTGACCTGATCCGACAGTACAGAGATGTTAGATATGATGA 583
QY 303 ATGCCGAACCATCCCGGAGGCTCCGAGAGAGATATGTGTAAACCAAAATGGGG 362
Db 584 ATGCCGAACCATCCCGGAGGCTCCGAGAGAGATATGTGTAAACCAAAATGGGG 643
QY 363 GATTTATGCAATTCCTCCGAGCAACCTGTGTATCGAGGGCCCTACTGSAACCCCTACTG 422
Db 644 GATTTATGCAATTCCTCCGAGCAACCTGTGTATCGAGGGCCCTACTGSAACCCCTACTG 703
QY 423 GACCCCTACTAGTCCGATACCCAGAGCTGCCCCACACACTGACGCTCCAAATATAC 482
Db 704 GACCCCTACTAGTCCGATACCCAGAGCTGCCCCACACACTGACGCTCCAAATATAC 763
QY 483 CACGATCTCAGAGGCTCTATATGCGGCTTTGGATACCAATGATGAAGCAACCAATG 542
Db 764 CACGATCTCAGAGGCTCTATATGCGGCTTTGGATACCAATGATGAAGCAACCAATG 823
QY 543 TGTGATGTGAGAGATGTCACAGATCCACACAGTGCACCCACCCAGATCTGCAT 602
Db 824 TGTGATGTGAGAGATGTCACAGATCCACACAGTGCACCCACCCAGATCTGCAT 883
QY 603 CAATCTGAAGCGGGGTACACCTGCTCTGACCCGAGGATATGGCTTGTGAAAGGCA 662
Db 884 CAATCTGAAGCGGGGTACACCTGCTCTGACCCGAGGATATGGCTTGTGAAAGGCA 943
QY 663 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTGTGGGAATGTTC 722
Db 944 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTGTGGGAATGTTC 1003
QY 723 TGAGTCTATTTCTTTGATACCAACCTGTGTTTACCCTCAATGAGATGGAAGGCTTGTG 782
Db 1004 TGAGTCTATTTCTTTGATACCAACCTGTGTTTACCCTCAATGAGATGGAAGGCTTGTG 1063
QY 783 CCAAGTGTGAGAGATGTGTCACACCGAAGACCTGTGCTGCAACCTGCTGACACCTA 842
Db 1064 CCAAGTGTGAGAGATGTGTCACACCGAAGACCTGTGCTGCAACCTGCTGCTGACACCTA 1123
QY 843 CGGCTCTTCATGTCGCGCTGTGACCCAGAGATGATGAATGAGAAAGATGGGCTTCATTTG 902
Db 1124 CGGCTCTTCATGTCGCGCTGTGACCCAGAGATGATGAATGAGAAAGATGGGCTTCATTTG 1183
QY 903 CAGTGAATGAGAGATGACGAGCTTCTGTGAGTTCTCTGACACATGAGTGTGAACCA 962
Db 1184 CAGTGAATGAGAGATGACGAGCTTCTGTGAGTTCTCTGACACATGAGTGTGAACCA 1243
QY 963 GCCCGGACATATCTTCTGCTCCCTCCAGGCTACATCTGCTGATGATGACACCGAAG 1022
Db 1244 GCCCGGACATATCTTCTGCTCCCTCCAGGCTACATCTGCTGATGATGACACCGAAG 1303
QY 1023 CTGCCAAGACATCAAGCAATGTGAGCAGAGAACACACGTCGAACCGCAGCAGACGTTG 1082
Db 1304 CTGCCAAGACATCAAGCAATGTGAGCAGAGAACACACGTCGAACCGCAGCAGACGTTG 1363
QY 1083 CTACAATTTACAAGGGGGCTTCAAAATGCATGACCCCAATCCGCTGTGAAGAGCCCTATCT 1142
Db 1364 CTACAATTTACAAGGGGGCTTCAAAATGCATGACCCCAATCCGCTGTGAAGAGCCCTATCT 1423
QY 1143 GAGGATCAGTATACCGCTGTATGTCTGCTGAGAGAACCTGTGCTGAGAGACGAC 1202
Db 1424 GAGGATCAGTATACCGCTGTATGTCTGCTGAGAGAACCTGTGCTGAGAGACGAC 1483
QY 1203 CTTTACATCTTGTACCGGACATGAGAGTGTGAGAGAGCTCCGTTCCGCTGACAT 1262
Db 1484 CTTTACATCTTGTACCGGACATGAGAGTGTGAGAGAGCTCCGTTCCGCTGACAT 1543
QY 1263 CTTTCAAAATGCAAGCAGACGACGCTACCTTGCGGCTATTTACATTTTCCAGATCAATC 1322
Db 1544 CTTTCAAAATGCAAGCAGACGACGCTACCTTGCGGCTATTTACATTTTCCAGATCAATC 1603
QY 1323 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAGCGGCCCATCACTGACCCCTGGT 1382
|||||

```

```

Db 1604 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAAACGGGCCCATCACTGACCCCTGGT 1663
QY 1383 GATGACACGGCCCATCAAAAGGCCCCGGGAATCCAGCTGATGSAATGATCACTGT 1442
Db 1664 GATGACACGGCCCATCAAAAGGCCCCGGGAATCCAGCTGATGSAATGATCACTGT 1723
QY 1443 CAACACTGTATCACTTACAGAGCAGCTCCGTGTATCCGACAGCGAGATATATGTGCGCA 1502
Db 1724 CAACACTGTATCACTTACAGAGCAGCTCCGTGTATCCGACAGCGAGATATATGTGCGCA 1783
QY 1503 GTACCATCTGTGACCTTCGGGCTGAGGCTCCAGCTGCTCTCATTTGGACCAAGGA 1562
Db 1784 GTACCATCTGTGACCTTCGGGCTGAGGCTCCGAGCTGCTCTCATTTGGACCAAGGA 1843
QY 1563 CAGAGAGAGAGAGAAATTAACAGAGATATAGAGCGACACAGAGATGAGGATTTCCG 1622
Db 1844 CAGAGAGAGAGAGAAATTAACAGAGATATAGAGCGACACAGAGATGAGGATTTCCG 1903
QY 1623 CTGAACGTTTCCCGGAGAGTCAAGCCCGACTTCTGACTCTACCTGTACTATTTGACA 1682
Db 1904 CTGAACGTTTCCCGGAGAGTCAAGCCCGACTTCTGACTCTACCTGTACTATTTGACA 1963
QY 1683 CCTGTACCTCTGCAGACTTGCACACCCCACTTCTATGATACATTTATCAAAAGATTT 1742
Db 1964 CCTGTACCTCTGCAGACTTGCACACCCCACTTCTATGATACATTTATCAAAAGATTT 2023
QY 1743 ATCATTCCTCCCTGATGAGAGATTTGTTGATGATTTTCAAGGCTTCACTTATTTCCA 1802
Db 2024 ATCATTCCTCCCTGATGAGAGATTTGTTGATGATTTTCAAGGCTTCACTTATTTCCA 2083
QY 1803 CTATTTTCAAGAAATATGATTTAGGTTTGGGGGCTGAGTCTATGTTCAAAAGCTGTG 1862
Db 2084 CTATTTTCAAGAAATATGATTTAGGTTTGGGGGCTGAGTCTATGTTCAAAAGCTGTG 2143
QY 1863 AAGAGCTTGTGACCTTCTTCACTCTTCCACCTCTCTGACATGCTTACTGCTTTG 1922
Db 2144 AAGAGCTTGTGACCTTCTTCACTCTTCCACCTCTCTGACATGCTTACTGCTTTG 2203
QY 1923 CAAAGACCCGAGAGCTGGCGGGAACCTGGGAGTAGTGTGCTTTTCCGTACACA 1982
Db 2204 CAAAGACCCGAGAGCTGGCGGGAACCTGGGAGTAGTGTGCTTTTCCGTACACA 2263
QY 1983 GAGAGGCTATGTAACAACAACACAGACAGATCGAAGGTTTTAGAGATGTGTTCAA 2042
Db 2264 GAGAGGCTATGTAACAACAACACAGACAGATCGAAGGTTTTAGAGATGTGTTCAA 2323
QY 2043 AACCATGCTGTGTTTTTCAACCAATTAAGAAAGTTCACTGTTTAAATTTGATTAAC 2102
Db 2324 AACCATGCTGTGTTTTTCAACCAATTAAGAAAGTTCACTGTTTAAATTTGATTAAC 2383
QY 2103 GGTTAATTTCTGTGTTTCAATTTTGAATTTTAAATAATATGTGAGATTCCTTC 2162
Db 2384 GGTTAATTTCTGTGTTTCAATTTTGAATTTTAAATAATATGTGAGATTCCTTC 2443
QY 2223 AGCCCAATGTTTTTTTGGAGACCCCTTAATCTGCTTCTTAAATTTTAAACCAAT 2282
Db 2444 GAAAGGCTTCAGACACATGCTATGTTCTGCTTCCAAACCAACCTGCTGCTCAATTT 2503
QY 2283 GGAATGGAATGAGAGTCTTCAACAGTATTAATATTTGAAGAGA 2328
Db 2564 GGAATGGAATGAGAGTCTTCAACAGTATTAATATTTGAAGAGA 2609

```

RESULT 13
 US-10-176-913-407
 ; Sequence 407, Application US/10176913
 ; Publication No. US20030022298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.

1184 CAGGATATGAGACAGTGCACCTTCTGAGTTCCTCGCCAACTGATGTGTACCA 1243
963 GCCCGGACATATCTTCTGCTCTGCTCCAGCTACATCTCTGTGATGACACCGAAG 1022
1244 GCCCGGACATATCTTCTGCTCTGCTCCAGCTACATCTCTGTGATGACACCGAAG 1303
1023 CTGCGAAGACATCAACGATGTGAGCAGACAGAACACACAGTGCACACCTGACAGACGNG 1082
1304 CTGCGAAGACATCAACGATGTGAGCAGACAGAACACACAGTGCACACCTGACAGACGNG 1363
1083 CTGCAATTTACAAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCT 1142
1364 CTGCAATTTACAAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCT 1423
1143 GAGGATCAGTATTAACGCTGTATGTCTCTGTAGAAACCTGCTGACAGACAGACG 1202
1424 GAGGATCAGTATTAACGCTGTATGTCTCTGTAGAAACCTGCTGACAGACAGACG 1483
1203 CTTTACCATTCTTGTACCGGACATGAGAGTGTGTACAGAGCTCCCTCCGCTGACAT 1262
1484 CTTTACCATTCTTGTACCGGACATGAGAGTGTGTGTACAGAGCTCCCTCCGCTGACAT 1543
1263 CTTTCCAAATGCAAGCCGACGACCCCTACCTCGGGGCTTATTCATTTTCCAGATCAATC 1322
1544 CTTTCCAAATGCAAGCCGACGACCCCTACCTCGGGGCTTATTCATTTTCCAGATCAATC 1603
1323 TGGGAATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCATTCAGTGGCACCCCTGT 1382
1604 TGGGAATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCATTCAGTGGCACCCCTGT 1663
1383 GATGACACGCCCCATCAAGAGGCCCCGGGAATTCAGCTGGACTTGAATATGATCACTGT 1442
1664 GATGACACGCCCCATCAAGAGGCCCCGGGAATTCAGCTGGACTTGAATATGATCACTGT 1723
1443 CAACACTGTCATCACTCTGAGAGGACCTCGTATCCGACTGGGATATATGTGTCCA 1502
1724 CAACACTGTCATCACTCTGAGAGGACCTCGTATCCGACTGGGATATATGTGTCCA 1783
1503 GTACCATTTCTGAGCCTCGGGCTGAGACCTCCGACGCTGCTCTCATTTGGACCAAGGGA 1562
1784 GTACCATTTCTGAGCCTCGGGCTGAGACCTCCGACGCTGCTCTCATTTGGACCAAGGGA 1843
1563 GAGGAAAGAGAGGAATTAACAGAGAAATGAGAGCAGACACAGCTTGGACATTTCTGTG 1622
1844 GAGGAAAGAGAGGAATTAACAGAGAAATGAGAGCAGACACAGCTTGGACATTTCTGTG 1903
1623 CTGAAGCTTTCCCGAAGATCAGCCCGACTTCTGATCTCTGACCTGATCTATTGAGA 1682
1904 CTGAAGCTTTCCCGAAGATCAGCCCGACTTCTGATCTCTGACCTGATCTATTGAGA 1963
1683 CCTGTACCCCTGAGAGCTTGGCCACCCAGTTCTATGATAGATTAATCAAAAGTATT 1742
1964 CCTGTACCCCTGAGAGCTTGGCCACCCAGTTCTATGATAGATTAATCAAAAGTATT 2023
1743 ATCATTTGCTCCCTGATAGAAATGTGTGTGAATTTTCAAGGCTTCACTTTATTCCA 1802
2024 ATCATTTGCTCCCTGATAGAAATGTGTGTGAATTTTCAAGGCTTCACTTTATTCCA 2083
1803 CTATTTTCAAAAGAAATAGATTAGTTTGTGCGGGGCTGAGTCTATGTTCAAAAGACTGTG 1862
2084 CTATTTTCAAAAGAAATAGATTAGTTTGTGCGGGGCTGAGTCTATGTTCAAAAGACTGTG 2143
1863 AACAGCTTGTCTCACTTCTTCACTCTTCACTCTTCTGACGTGTGTACTGCTTGTG 1922
2144 AACAGCTTGTCTCACTTCTTCACTCTTCTGACGTGTGTGTACTGCTTGTG 2203
1923 CAAAGACCCGGAGCTGGGGGGAACCCGAGAGTGTGAGTGTGTTTGTGTTTGGTACACA 1982
2204 CAAAGACCCGGAGCTGGGGGGAACCCGAGAGTGTGAGTGTGTTTGTGTTTGGTACACA 2263
1983 GAGAGGCTATGTAAACAAACACAGAGATGGAAGGCTTTTAAAGATGTGTTTCA 2042

2264 GAGAGGCTATGTAAACAAACACAGAGATGCAAGGCTTTTAAAGATGTGTTTCA 2323
2043 AACCATCCCTGTATTTTCAACATTAAGAAAGTTTCACTGTCTTAAATTTGTATAC 2102
2324 AACCATCCCTGTATTTTCAACATTAAGAAAGTTTCACTGTCTTAAATTTGTATAC 2383
2103 GGTATATCTGTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2162
2384 GGTATATCTGTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2443
2163 GAAAGGCTTCAAGACATGATGTGTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2222
2444 GAAAGGCTTCAAGACATGATGTGTCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2503
2223 AGCCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2282
2504 AGCCAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2563
2283 GATGTGAATGACAGAGTCTCCAAACGATTAATTTTGAAGAGA 2328
2564 GATGTGAATGACAGAGTCTCCAAACGATTAATTTTGAAGAGA 2609

RESULT 12
US-10-176-757-407
; Sequence 407, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R186
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO. 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

3 CCGGCGCTCTCCCGCTGCTCTCTCCAGCAGCTCGTGGGCGCTTGGAAATAAACACCC 62
284 CCGGCGCTCTCCCGCTGCTCTCTCCAGCAGCTCGTGGGCGCTTGGAAATAAACACCC 343
63 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGCGGAGCTCTCCGCGGCTCCGCGCCG 122
344 GCGAGCCCGGAGGCGCCAGAGAGGCGGAGCGGAGCTCTCCGCGGCTCCGCGCGG 403
123 CGAGCTTCTCTGCGCTTCTGATCTCTCTCTGCGGCGCTTGGAGATGCGAGAAATAA 182
404 CGAGCTTCTCTGCGCTTCTGATCTCTCTCTGCGGCGCTTGGAGATGCGAGAAATAA 463
183 AAGATATCACTGTATACATTTCTGCTCTCTTCAAGAGCCCGGGAATGCAAGGC 242
464 AAGATATCACTGTATACATTTCTGCTCTCTTCAAGAGCCCGGGAATGCAAGGC 523
243 AAGATATCACTGTATACATTTCTGCTCTCTTCAAGAGCCCGGGAATGCAAGGC 302

Db 2504 AGCCAGGCTTTTCCTTTAGGAGCCCTTAATCTGCTTCTTTAGAAATTTTACCCAAAT 2563

Qy 2283 GGATTGGAAATGCAGAGCTCTCCAAACTGATTAATAATTTGAAGAGA 2328

Db 2564 GGATTTGGAATGCAGAGCTCTCCAAACTGATTAATAATTTGAAGAGA 2609

RESULT 10

US-10-175-752-407
; Sequence 407, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION.

[illegible]

	Matches	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	3	CCGGGACCTCTCCCGGTCCTCTCCACAGCATCGCTGGGGCCCTGTGAAATMAACACC	62						
Db	284	CCGGGACCTCTCCCGGTCCTCTCCACAGCATCGCTGGGGCCCTGTGAAATMAACACC	343						
QY	63	GCGAGCCCGGAGGCGCCAGAGGAGCGACGTGCCGAGCTCTCCGGGGGTCGGCCCG	122						
Db	344	GCGAGCCCGGAGGCGCCAGAGGAGCGACGTGCCGAGCTCTCCGGGGGTCGGCCCG	403						
QY	123	CGAGCTTCTCTTCGCGCTCGCATCTCTCTCGCGGCTCTTGACATGCCAGAAATAA	182						
Db	404	CGAGCTTCTCTTCGCGCTCGCATCTCTCTCGCGGCTCTTGACATGCCAGAAATAA	463						
QY	183	AAGGATACACTGTTACCAATTCGGGCTCTGTCTTCCAAAGCCCTGGGAATGCACAGC	242						
Db	464	AAGGATACACTGTTACCAATTCGGGCTCTGTCTTCCAAAGCCCTGGGAATGCACAGC	523						
QY	243	ACAGTGCACCAATGCTTTAGACTGTATGCGCACTAGAGACAGTGTATATTGATGA	302						
Db	524	ACAGTGCACCAATGCTTTAGACTGTATGCGCACTAGAGACAGTGTATATTGATGA	583						
QY	303	ATACCGAACCATCCCGAGGCTGCCGAGAGACATATATGTGTTTAACCAAAATGGCGG	362						
Db	584	ATACCGAACCATCCCGAGGCTGCCGAGAGACATATATGTGTTTAACCAAAATGGCGG	643						
QY	363	GTATTATGCAATCCCGGACAAACCTGTATCGAGGGGCCCTACGTGAACCCCTACTC	422						
Db	644	GTATTATGCAATCCCGGACAAACCTGTGTATCGAGGGGCCCTACTGTGAACCCCTACTC	703						
QY	423	GACCCCTACTCAGGTCCGTACCAAGCAGCTGCCCAACACTCTCAGCTCCAAATAATCC	482						
Db	704	GACCCCTACTCAGGTCCGTACCAAGCAGCTGCCCAACACTCTCAGCTCCAAATAATCC	763						

Oy 723 TGGATCTATCTGTACATGCAACCGTGGTTTACCTCAATGAGATGAGAGGCTTG 782
 Db 1004 TGGATCTATCTGTACATGCAACCGTGGTTTACCTCAATGAGATGAGAGGCTTG 1063
 Oy 783 CCAAGATGTAAGAGAGTGTGCCACGAGAACCCCTGGCTGCAAACTGCTCAACACCTA 842
 Db 1064 CCAAGATGTAAGAGAGTGTGCCACGAGAACCCCTGGCTGCAAACTGCTCAACACCTA 1123
 Oy 843 CGGCTCTTCACTGTCCCTGTGACCCAGATATGAACCTTGAGAGAGATGGCTTCAATG 902
 Db 1124 CGGCTCTTCACTGTCCCTGTGACCCAGATATGAACCTTGAGAGAGATGGCTTCAATG 1183
 Oy 903 CAGTATATGAGAGAGTGTGACCTTCTGTGATTCCTTGGCAACATGAGTGTGAACA 962
 Db 1184 CAGTATATGAGAGAGTGTGACCTTCTGTGATTCCTTGGCAACATGAGTGTGAACA 1243
 Oy 963 GCCCGGACATCTTCTGCTCCCTGACCTCCAGGCTACATCTCTGTGATGACCAACGAG 1022
 Db 1244 GCCCGGACATCTTCTGCTCCCTGACCTCCAGGCTACATCTCTGTGATGACCAACGAG 1303
 Oy 1023 CTGCCAAGCATCAACGATGTGACAGCAAGAACACAGTGCACCTGACAGACAGCTG 1082
 Db 1304 CTGCCAAGCATCAACGATGTGACAGCAAGAACACAGTGCACCTGACAGACAGCTG 1363
 Oy 1083 CTACAATTTACAAGGGGGCTTCAAAATGATCGAACCCATCGCTGTGAGAGAGCTTATCT 1142
 Db 1364 CTACAATTTACAAGGGGGCTTCAAAATGATCGAACCCATCGCTGTGAGAGAGCTTATCT 1423
 Oy 1143 GAGGATCACTGATTAACCGCTGTATGTCTCTGTGAGAACCCCTGGCTGACAGACAGCC 1202
 Db 1424 GAGGATCACTGATTAACCGCTGTATGTCTCTGTGAGAACCCCTGGCTGACAGACAGCC 1483
 Oy 1203 CTTTACCATCTGTACCGGGGACATGAGACGCTGTGAGACGCTCCCTTCCCGCTGACAT 1262
 Db 1484 CTTTACCATCTGTGTACCGGGGACATGAGACGCTGTGAGACGCTCCCTTCCCGCTGACAT 1543
 Oy 1263 CTTTCCAAATGCAAGCCACGACCCGCTACCCCTGAGGCTATTTACATTTTCCAGATCAATC 1322
 Db 1544 CTTTCCAAATGCAAGCCACGACCCGCTACCCCTGAGGCTATTTACATTTTCCAGATCAATC 1603
 Oy 1323 TEGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGGCCCATGAGTGCACCTGCT 1382
 Db 1604 TEGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGGCCCATGAGTGCACCTGCT 1663
 Oy 1383 GATGACAGGGGGCCATCAAGGGGGGGGAAATCCAGCTGAGATTTGAAAATGATCACTGT 1442
 Db 1664 GATGACAGGGGGCCATCAAGGGGGGGGAAATCCAGCTGAGATTTGAAAATGATCACTGT 1723
 Oy 1443 CAACACTGTCACTCACTTCAAGAGGAGCTCCGATCCGACTGCGGATATATGTGTGCGA 1502
 Db 1724 CAACACTGTCACTCACTTCAAGAGGAGCTCCGATCCGACTGCGGATATATGTGTGCGA 1783
 Oy 1503 GTACCACTTCTGAGCCTCGGGCTGAGAGCTCCGACGCTGCTCTCATTTGACCAAGGGA 1562
 Db 1784 GTACCACTTCTGAGCCTCGGGCTGAGAGCTCCGACGCTGCTCTCATTTGACCAAGGGA 1843
 Oy 1563 CAGGAGAGAGAGAGAAATAACAGAGAGAAATGAGAGGACACAGAGCTTAGGCAATTCCTG 1622
 Db 1844 CAGGAGAGAGAGAGAAATAACAGAGAGAAATGAGAGGACACAGAGCTTAGGCAATTCCTG 1903
 Oy 1623 CTGAACGTTTCCCGGAAAGATCAGCCCGACTTCTGACTCTCACTGTAATTTGACAGA 1682
 Db 1904 CTGAACGTTTCCCGGAAAGATCAGCCCGACTTCTGACTCTCACTGTAATTTGACAGA 1963
 Oy 1683 CCTGTACCCCTGACAGAGCTTCCACCCCAAGTTCTATGATACAGTTATCAAAAAGTATT 1742
 Db 1964 CCTGTACCCCTGACAGAGCTTCCACCCCAAGTTCTATGATACAGTTATCAAAAAGTATT 2023
 Oy 1743 ATCATTTGCTCCCTGATTAAGATTTGTGTGTAATTTTCAAGGCTTCAAGTTATTCCA 1802
 Db 2024 ATCATTTGCTCCCTGATTAAGATTTGTGTGTAATTTTCAAGGCTTCAAGTTATTCCA 2083

Oy 1803 CTAATTTCAAGAAATAGATTAGTTGCGGGGCTGAGCTATGTTCAAGACGTGG 1862
 Db 2084 CTAATTTCAAGAAATAGATTAGTTGCGGGGCTGAGCTATGTTCAAGACGTGG 2143
 Oy 1863 AACAGCTTGTCTACTTCTTCACTCTTCCACTCTCTCTCACTGTGTTACTGCTTG 1922
 Db 2144 AACAGCTTGTCTACTTCTTCACTCTTCCACTCTCTCTCACTGTGTTACTGCTTG 2203
 Oy 1923 CAAGAGCCGGGAGCTGGCGGGGAAACCTGGGAGTACTGATTTGCTTTGCGTACCA 1982
 Db 2204 CAAGAGCCGGGAGCTGGCGGGGAAACCTGGGAGTACTGATTTGCTTTGCGTACCA 2263
 Oy 1983 GAGAGGCTATGTAACAAACACAGAGATCCAGAGGTTTGAAGATGTTTCA 2042
 Db 2264 GAGAGGCTATGTAACAAACACAGAGATCCAGAGGTTTGAAGATGTTTCA 2323
 Oy 2043 AACCATGCTGTATTTTCAACATTAAGAGATTTCAAGTGTCTTAAATTTGTAAC 2102
 Db 2324 AACCATGCTGTATTTTCAACATTAAGAGATTTCAAGTGTCTTAAATTTGTAAC 2383
 Oy 2103 GGTATATCTGTCTGTCTGATTTGATTTTAAATAATGCTAGATTCCTTC 2162
 Db 2384 GGTATATCTGTCTGTCTGATTTGATTTTAAATAATGCTAGATTCCTTC 2443
 Oy 2163 GAAAGCCTTCAGACATGCTATGTTCTGTCTTCCAAACAGTCTCTCCATTTT 2222
 Db 2444 GAAAGCCTTCAGACATGCTATGTTCTGTCTTCCAAACAGTCTCTCCATTTT 2503
 Oy 2223 AGCCAGGTGTTTCTTTGAGAGCCCTTAATCTTGTCTTTTGAATTTTACCATT 2282
 Db 2504 AGCCAGGTGTTTCTTTGAGAGCCCTTAATCTTGTCTTTTGAATTTTACCATT 2563
 Oy 2283 GGAATGAATGACAGAGGTCTCCAAACGATTAATTTTGAAGGA 2328
 Db 2564 GGAATGAATGACAGAGGTCTCCAAACGATTAATTTTGAAGGA 2609

RESULT 9

US-10-175-738-407

; Sequence 407, Application US/10175738
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Collin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C45
 ; CURRENT APPLICATION NUMBER: US/10/175,738
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior application removed - See file wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 407
 ; LENGTH: 2609
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-175-738-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 3 CCGGCGCTCTCCCGCTGCTCTCCACGAGCTGCTGGCCCTCTGTGAATTAACACCC 62
 Db 284 CCGGCGCTCTCCCGCTGCTCTCCACGAGCTGCTGGCCCTCTGTGAATTAACACCC 343

```

QY 1383 GATGACAGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGGACTGGAAATGATCAGCTGT 1442
D 1664 GATGACAGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGGACTGGAAATGATCAGCTGT 1723
QY 1443 CAACACTGTCATCACTTCAGAGGCGAGCTCGGTGATCCGACTGCGGATATATGTGTGCGA 1502
D 1724 CAACACTGTCATCACTTCAGAGGCGAGCTCGGTGATCCGACTGCGGATATATGTGTGCGA 1783
QY 1503 GTACCCATTCGAGCCCTGGGCTGGAGCTCCGAGCCGCTCCATTCGATGGCCACAGGGA 1562
D 1784 GTACCCATTCGAGCCCTGGGCTGGAGCTCCGAGCCGCTCCATTCGATGGCCACAGGGA 1843
QY 1563 CAGGAGAGAGAGAGAAATTAACAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
D 1844 CAGGAGAGAGAGAGAAATTAACAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1903
QY 1623 CTGAGCTTTCCCGGAGAGAGTCCAGCCCGAGCTCTGACTCTGACTCTGACTCTGACTCTG 1682
D 1904 CTGAGCTTTCCCGGAGAGAGTCCAGCCCGAGCTCTGACTCTGACTCTGACTCTGACTCTG 1963
QY 1683 CTGTCACCTGCGAGAGTCCAGCCCGAGCTCTGACTCTGACTCTGACTCTGACTCTGACTCTG 1742
D 1964 CTGTCACCTGCGAGAGTCCAGCCCGAGCTCTGACTCTGACTCTGACTCTGACTCTGACTCTG 2023
QY 1743 ATCATTGCTCCCTGATAGAGATGTTGTTGTAATTTTCAAGGCTTCAGTTTATTTCCA 1802
D 2024 ATCATTGCTCCCTGATAGAGATGTTGTTGTAATTTTCAAGGCTTCAGTTTATTTCCA 2083
QY 1803 CTATTTTCAAGAAATAGATTAGATTGGGGGCTGAGTCTATGTTCAAAAGACTGTG 1862
D 2084 CTATTTTCAAGAAATAGATTAGATTGGGGGCTGAGTCTATGTTCAAAAGACTGTG 2143
QY 1863 AACAGCTGCTGCTACTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1922
D 2144 AACAGCTGCTGCTACTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2203
QY 1923 CAAGAGCCCGGAGAGCTGGCGGGAGAACCTGGGAGTAGTCTGTTTCTCTCTCTCTCTCTCT 1982
D 2204 CAAGAGCCCGGAGAGCTGGCGGGAGAACCTGGGAGTAGTCTGTTTCTCTCTCTCTCTCTCT 2263
QY 1983 GAGAGAGCTATGTAACAAACACAGAGATCCAGAGGTTTTCAGAGATGTTTCA 2042
D 2264 GAGAGAGCTATGTAACAAACACAGAGATCCAGAGGTTTTCAGAGATGTTTCA 2323
QY 2043 AACCATGCTGCTGTTTTCACATTAACATTAAGAGTTCAGTGTCTCTTAATTTGATTAAC 2102
D 2324 AACCATGCTGCTGTTTTCACATTAACATTAAGAGTTCAGTGTCTCTTAATTTGATTAAC 2383
QY 2103 GGTTTAATCTGCTGTTTTCATTTTGAATTTTAAAAAATATGTCGTAATTCCTTC 2162
D 2384 GGTTTAATCTGCTGTTTTCATTTTGAATTTTAAAAAATATGTCGTAATTCCTTC 2443
QY 2163 GAAAGGCTTCAGACATGTAATGTTCTGTTTCCAAACCACTCTCTCTCTCTCTCTCTCT 2222
D 2444 GAAAGGCTTCAGACATGTAATGTTCTGTTTCCAAACCACTCTCTCTCTCTCTCTCTCT 2503
QY 2223 AGCCCACTGTTTCTTGGAGAGCCCTTAATCTGTTTCTTGAATTTTAAATTTTAAACCAAT 2282
D 2504 AGCCCACTGTTTCTTGGAGAGCCCTTAATCTGTTTCTTGAATTTTAAATTTTAAACCAAT 2563
QY 2283 GATTTGGAATGACAGAGTCTCCAAATGATTAATTTTGAAGAGA 2328
D 2564 GATTTGGAATGACAGAGTCTCCAAATGATTAATTTTGAAGAGA 2609

```

```

RESULT 8
US-10-173-706-407
; Sequence 407, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan

```

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenli
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; PRIOR APPLICATION: 2002-06-17
; PRIOR APPLICATION: 2002-06-17
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-407

Query Match 99.8% Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9% Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGGGGCTCTCCCGCTGCTCTCTCCAGACTCGCTCGGCCCCCTGGAATAAACACCC 62
D 284 CCGGGGCTCTCCCGCTGCTCTCTCCAGACTCGCTCGGCCCCCTGGAATAAACACCC 343
QY 63 GCGAGCCCGAGGGCCCGAGAGAGCGAGCTGCCGAGCTCTCGGGGCTCCCGCCG 122
D 344 GCGAGCCCGAGGGCCCGAGAGAGCGAGCTGCCGAGCTCTCGGGGCTCCCGCCG 403
QY 123 CGAGCTTTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
D 404 CGAGCTTTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 463
QY 183 AAGATATCTACTGTTTACATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
D 464 AAGATATCTACTGTTTACATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
QY 243 ACACTGACGAGATGCTTGAACCTGATGCGCCAGTCCAGAGAGAGTGTGATATGATGA 302
D 524 ACACTGACGAGATGCTTGAACCTGATGCGCCAGTCCAGAGAGAGTGTGATATGATGA 583
QY 303 ATGCGGAACATCCCCGAGGCTCGCGAGAGAGACATGATGTGTTAAACAAATGGCGG 362
D 584 ATGCGGAACATCCCCGAGGCTCGCGAGAGAGACATGATGTGTTAAACAAATGGCGG 643
QY 363 GTATTATGCAATTCCTCGGAGAAACCTGTGTATGAGAGGCTCTACTCGAACCCCTACTC 422
D 644 GTATTATGCAATTCCTCGGAGAAACCTGTGTATGAGAGGCTCTACTCGAACCCCTACTC 703
QY 423 GACCCCTACTCAGGTCCGTAACCAAGAGCTGCCCCACACTCTCAGTCCAAACTATCC 482
D 704 GACCCCTACTCAGGTCCGTAACCAAGAGCTGCCCCACACTCTCAGTCCAAACTATCC 763
QY 483 CAGCATCTCCAGGCTCTTATATGCGCTTGGATGATACAGATGATGAAGAACCAAG 542
D 764 CAGCATCTCCAGGCTCTTATATGCGCTTGGATGATACAGATGATGAAGAACCAAG 823
QY 543 TGTGATGTGAGAGAGTGTGCAACAGATTCACAGAGTGCACACCCACAGATCGCAT 602
D 824 TGTGATGTGAGAGAGTGTGCAACAGATTCACAGAGTGCACACCCACAGATCGCAT 883
QY 603 CAATATCTGAAGGCGGGTACACTGCTCTGACAGCGAGAGATTTGGCTTCTGGAAGCCA 662
D 884 CAATATCTGAAGGCGGGTACACTGCTCTGACAGCGAGATTTGGCTTCTGGAAGCCA 943
QY 663 GTGCTTACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
D 944 GTGCTTACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1003

```

QY 2043 AACCATGCCGTGATTTTCAACCATTAAGAGATTTCAGTGTCTTAATTTGATTAAC 2102
| | | | |
Db 2324 AACCATGCCGTGATTTTCAACCATTAAGAGATTTCAGTGTCTTAATTTGATTAAC 2383
| | | | |
QY 2103 GGTTAATTCGTCTGTTGATTTTGAATTTTAAAAAATATGCTAGAAATTCCTTC 2162
| | | | |
Db 2384 GGTTAATTCGTCTGTTGATTTTGAATTTTAAAAAATATGCTAGAAATTCCTTC 2443
| | | | |
QY 2163 GAAAGGCTTCACACACATGCTATGCTGCTTCCCAACCCAGTCTCTCCATTTT 2222
| | | | |
Db 2444 GAAAGGCTTCACACACATGCTATGCTGCTTCCCAACCCAGTCTCTCCATTTT 2503
| | | | |
QY 2223 AGCCAGTGTCTTCTTGAGAGACCCCTTAATCTGCTTTTGAATTTTAAACCAAT 2282
| | | | |
Db 2504 AGCCAGTGTCTTCTTGAGAGACCCCTTAATCTGCTTTTGAATTTTAAACCAAT 2563
| | | | |
QY 2283 GGATTGGAATGACAGAGTCTCAAACTGATTAATTTGAAGAGA 2328
| | | | |
b 2564 GGATTGGAATGACAGAGTCTCAAACTGATTAATTTGAAGAGA 2609
| | | | |
RESULT 7
US-10-175-737-407
; Sequence 407, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Collin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Saplen
US-10-175-737-407
Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CCCGGCGCTCTCCCGTGTCTCTCCACGACTGCGCTCGCCCTCTGGAATTAACAACCC 62
| | | | |
Db 284 CCCGGCGCTCTCCCGTGTCTCTCCACGACTGCGCTCGCCCTCTGGAATTAACAACCC 343
| | | | |
QY 63 GCGACCCCGAGAGGCCCAAGAGAGCGAGCTGCCGAGCTCTCGGGGGTCCCGCCG 122
| | | | |
Db 344 GCGACCCCGAGAGGCCCAAGAGAGCGAGCTGCCGAGCTCTCGGGGGTCCCGCCG 403
| | | | |
QY 123 CGAGCTTCTTCTGCGCTCGCATCTCTCCGCGGCTTGGACATGGCAGAAATAA 182
| | | | |
Db 404 CGAGCTTCTTCTGCGCTCGCATCTCTCCGCGGCTTGGACATGGCAGAAATAA 463
| | | | |
QY 183 AAGGATACCTACTGTTACCATTTCTGCTCTCTTCCAAAGCCCTGGAAATGACAGGC 242
| | | | |
Db 464 AAGGATACCTACTGTTACCATTTCTGCTCTCTTCCAAAGCCCTGGAAATGACAGGC 523
| | | | |
QY 243 AAGGATACCTACTGTTACCATTTCTGCTCTCTTCCAAAGCCCTGGAAATGACAGGC 302
| | | | |

Db 524 ACAGTGACGAGATGGCTTTGACCTGATCGCCAGTCCAGAGAGTGTATGATATGATGA 583
| | | | |
QY 303 ATCCCGAACCATCCCCGAGGCTCCGAGAGAGATGATGTGTAAACCAAAATGCGGG 362
| | | | |
Db 584 ATCCCGAACCATCCCCGAGGCTCCGAGAGAGATGATGTGTAAACCAAAATGCGGG 643
| | | | |
QY 363 GATTTATGATTTCCCGGGAACCAACCTGTGTATCGAGGGGCCCTACTGAAACCCCTACTC 422
| | | | |
Db 644 GATTTATGATTTCCCGGGAACCAACCTGTGTATCGAGGGGCCCTACTGAAACCCCTACTC 703
| | | | |
QY 423 GACCCCTACTCAGTCCGTATCCAGAGAGTGGCCCAACACTCAGTCCAAATATCC 482
| | | | |
Db 704 GACCCCTACTCAGTCCGTATCCAGAGAGTGGCCCAACACTCAGTCCAAATATCC 763
| | | | |
QY 483 CACGATCTCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAACCAACATG 542
| | | | |
Db 764 CACGATCTCAGGCTCTTATATGCGCTTTGGATACAGATGATGAAACCAACATG 823
| | | | |
QY 543 TGTGATGTGAGAGAGTGTGCAACAGATTGCCACCACTGCAACCCCAACCATGAT 602
| | | | |
Db 824 TGTGATGTGAGAGAGTGTGCAACAGATTGCCACCACTGCAACCCCAACCATGAT 883
| | | | |
QY 603 CAATACGAGAGGCGGTACACCTGCTCTGACCGAGAGATATGCTTGTGAAGGCCA 662
| | | | |
Db 884 CAATACGAGAGGCGGTACACCTGCTCTGACCGAGAGATATGCTTGTGAAGGCCA 943
| | | | |
QY 663 GTGCTTAGACATGATGATGTGCTATGCTTACTGCCAGAGCTGTGCGAAATGTTC 722
| | | | |
Db 944 GTGCTTAGACATGATGATGTGCTATGCTTACTGCCAGAGCTGTGCGAAATGTTC 1003
| | | | |
QY 723 TGGATCTTATCTTGTATACATGCAACCTGTTTAACTCAATAGATGGAAGTCTTG 782
| | | | |
Db 1004 TGGATCTTATCTTGTATACATGCAACCTGTTTAACTCAATAGATGGAAGTCTTG 1063
| | | | |
QY 783 CCAAGATGTGAACGAGTGTGCAACCGAGAACCCCTGCGTCAAACTGCGTCAACACTA 842
| | | | |
Db 1064 CCAAGATGTGAACGAGTGTGCAACCGAGAACCCCTGCGTCAAACTGCGTCAACACTA 1123
| | | | |
QY 843 CGGCTTCTTATCTGCGGCTGTGACCAAGATATGAACCTGAGAGATGCGCTTCAATG 902
| | | | |
Db 1124 CGGCTTCTTATCTGCGGCTGTGACCAAGATATGAACCTGAGAGATGCGCTTCAATG 1183
| | | | |
QY 903 CAGTGTATGAGAGAGTGTGCAACCTGCTGCTGCAACATGATGATGGAACCA 962
| | | | |
Db 1184 CAGTGTATGAGAGAGTGTGCAACCTGCTGCTGCAACATGATGATGGAACCA 1243
| | | | |
QY 963 GCCCGGACATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGATGACAAACCGAAG 1022
| | | | |
Db 1244 GCCCGGACATCTTCTGCTCTGCGCTCCAGGCTACATCTGCTGATGACAAACCGAAG 1303
| | | | |
QY 1023 CTGCGCAAGATCAACGATGTGACAGAGAAACACACGCTGCACCTGACAGAGCTG 1082
| | | | |
Db 1304 CTGCGCAAGATCAACGATGTGACAGAGAAACACACGCTGCACCTGACAGAGCTG 1363
| | | | |
QY 1083 CTACAAATTAAGAGGGGCTTCAATGATCAACGACCCATCCGCTGTAGAGAGCTTATCT 1142
| | | | |
Db 1364 CTACAAATTAAGAGGGGCTTCAATGATCAACGACCCATCCGCTGTAGAGAGCTTATCT 1423
| | | | |
QY 1143 GAGGATCACTGATTAACCGCTGATGTCTCTGTGAAACCTGCTGACAGAGACAGCC 1202
| | | | |
Db 1424 GAGGATCACTGATTAACCGCTGATGTCTCTGTGAAACCTGCTGACAGAGACAGCC 1483
| | | | |
QY 1203 CTTTACCATCTGTATACCGGAGACATGAGACGTGTGAGAGAGCTTCCGCTGACAT 1262
| | | | |
Db 1484 CTTTACCATCTGTATACCGGAGACATGAGACGTGTGAGAGAGCTTCCGCTGACAT 1543
| | | | |
QY 1263 CTTTCAAAATGCAAGCAGACAGCCGCTAACCTGCGGCTTATTAATTTCCAGATCAATC 1322
| | | | |
Db 1544 CTTTCAAAATGCAAGCAGACAGCCGCTAACCTGCGGCTTATTAATTTCCAGATCAATC 1603
| | | | |
QY 1323 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACAGGGGCCCATCATGATGCCACCTGTG 1382
| | | | |
Db 1604 TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACAGGGGCCCATCATGATGCCACCTGTG 1663
| | | | |

NUMBER OF SEQ ID NOS: 612
 : SEQ ID NO 407
 : LENGTH: 2609
 : TYPE: DNA
 : ORGANISM: Homo Sapien
 : US-10-176-758-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGGGGCTTCCCGTCTCTCCACGACTGCTGGCCCTCTGGAAATAAACCC 62
 DB 284 CCGGGGCTTCCCGTCTCTCCACGACTGCTGGCCCTCTGGAAATAAACCC 343
 QY 63 GCGAGCCCGAGGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 122
 DB 344 GCGAGCCCGAGGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 403
 QY 123 CGAGCTTCTCTGCGCTTGGCATCTCTCTCGGCGCTCTGGACATGCCAGATAA 182
 DB 404 CGAGCTTCTCTGCGCTTGGCATCTCTCTCGGCGCTCTGGACATGCCAGATAA 463
 QY 183 AAGGATCTACTGTTACCATTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
 DB 464 AAGGATCTACTGTTACCATTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
 QY 243 ACAGTCAGAGATGCTTTGACCTGGATCGCCAGTCAGAGAGAGAGAGAGAGAG 302
 DB 524 ACAGTCAGAGATGCTTTGACCTGGATCGCCAGTCAGAGAGAGAGAGAGAGAG 583
 QY 303 ATGCCGACCATCCCGAGGCTCCGAGAGAGATGTTGTTAAACCAAAATGGCGG 362
 DB 584 ATGCCGACCATCCCGAGGCTCCGAGAGAGATGTTGTTAAACCAAAATGGCGG 643
 QY 363 GTATTATGATTTCCCGGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTCT 422
 DB 644 GTATTATGATTTCCCGGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTCT 703
 QY 423 GACCCCTACTAGTCCGTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
 DB 704 GACCCCTACTAGTCCGTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763
 QY 483 CACGATCTCCAGGCTCTTATATGCCCTTTGGATACAGATGATGAAGAACCAAT 542
 DB 764 CACGATCTCCAGGCTCTTATATGCCCTTTGGATACAGATGATGAAGAACCAAT 823
 QY 543 TGTGATGTGAGAGAGTGTGAACAGATTCACACAGTCCAAACCCCAATCTGCAT 602
 DB 824 TGTGATGTGAGAGAGTGTGAACAGATTCACACAGTCCAAACCCCAATCTGCAT 883
 QY 603 CAATTAAGAGGGGTGACACCTGCTCTGACAGCAGAGAGATTTGGCTTCTGGAAGGCA 662
 DB 884 CAATTAAGAGGGGTGACACCTGCTCTGACAGCAGAGAGATTTGGCTTCTGGAAGGCA 943
 QY 663 GTCTTAAGACATGATGATGCTATGTTACTGACAGCAGCTCTGTGCAATGTTCC 722
 DB 944 GTCTTAAGACATGATGATGCTATGTTACTGACAGCAGCTCTGTGCAATGTTCC 1003
 QY 723 TGTATCTATTTCTTGTACATGCAACCTGTTTAACTTCAATGAGATGGAAGCTTG 782
 DB 1004 TGTATCTATTTCTTGTACATGCAACCTGTTTAACTTCAATGAGATGGAAGCTTG 1063
 QY 783 CCAAGATGTGAAGAGTGTCCACAGAGAAACCCCTGCTGCAAACTGAGTCAACCTA 842
 DB 1064 CCAAGATGTGAAGAGTGTCCACAGAGAAACCCCTGCTGCAAACTGAGTCAACCTA 1123
 QY 843 CGGCTCTTCACTGCGCTGTGTGACCAAGATATGAATGAAGAGATGGCTTCAATG 902
 DB 1124 CGGCTCTTCACTGCGCTGTGTGACCAAGATATGAATGAAGAGATGGCTTCAATG 1183
 QY 903 CAGTGATGAGAGAGAGAGCTTCTGAGATTTCTTCCCAACATGAGTGTGACCA 962
 DB 1184 CAGTGATGAGAGAGAGAGCTTCTGAGATTTCTTCCCAACATGAGTGTGACCA 1243

DB 1184 CAGTGATGAGAGAGAGAGCTTCTGAGATTTCTTCCCAACATGAGTGTGACCA 1243
 QY 963 GCCGGGACATATTCTGCTCTGCGCCCTCCAGGCTACATCTCGTATGATGACACGAG 1022
 DB 1244 GCCGGGACATATTCTGCTCTGCGCCCTCCAGGCTACATCTCGTATGATGACACGAG 1303
 QY 1023 CTGGCAAGACATCAACGATGTGACACAGAGAAACCAAGTGCACCTGACAGACCTG 1082
 DB 1304 CTGGCAAGACATCAACGATGTGACACAGAGAAACCAAGTGCACCTGACAGACCTG 1363
 QY 1083 CTACAAATTTACAAAGGGGCTTCAATGATGACACCCATCCGCTGTGAGAGACCTTATCT 1142
 DB 1364 CTACAAATTTACAAAGGGGCTTCAATGATGACACCCATCCGCTGTGAGAGACCTTATCT 1423
 QY 1143 GAGATGATGATTAACCGCTGATGTGCTCGCTGAGAAACCTGGCTGACAGACAGCC 1202
 DB 1424 GAGATGATGATTAACCGCTGATGTGCTCGCTGAGAAACCTGGCTGACAGACAGCC 1483
 QY 1203 CTTTACCATCTTGTACCGGACATGAGAGCTGTGTGACAGACCTCTGTTCCGCTGACAT 1262
 DB 1484 CTTTACCATCTTGTACCGGACATGAGAGCTGTGTGACAGACCTCTGTTCCGCTGACAT 1543
 QY 1263 CTTTCAATATGCAAGCAGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAAAATC 1322
 DB 1544 CTTTCAATATGCAAGCAGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAAAATC 1603
 QY 1323 TGGGAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGCCCATCATGAGCCACCTGGT 1382
 DB 1604 TGGGAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGCCCATCATGAGCCACCTGGT 1663
 QY 1383 GATGACAGCCCATCAATCAAAAGGGCCCGGAAATCCAGCTGACTTGGAAATGATCACTGT 1442
 DB 1664 GATGACAGCCCATCAATCAAAAGGGCCCGGAAATCCAGCTGACTTGGAAATGATCACTGT 1723
 QY 1443 CAACACTGATCAACTTCAAGAGAGAGTCCGCTGATCCGACTGCGGATATATGTGCGCA 1502
 DB 1724 CAACACTGATCAACTTCAAGAGAGAGTCCGCTGATCCGACTGCGGATATATGTGCGCA 1783
 QY 1503 GTACCCATTTGAGACCTGCGGCTGAGGCTCCGAGCTGCTCTCTATTTGGCAACCAAGGA 1562
 DB 1784 GTACCCATTTGAGACCTGCGGCTGAGGCTCCGAGCTGCTCTCTATTTGGCAACCAAGGA 1843
 QY 1563 CAGGAGAGAGAGAGAAATTAACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1622
 DB 1844 CAGGAGAGAGAGAGAAATTAACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1903
 QY 1623 CTGAGACCTTCCCGAAGAGTCAAGCCGAGCTTCTGACTCTCAGCTGACTGATGACGA 1682
 DB 1904 CTGAGACCTTCCCGAAGAGTCAAGCCGAGCTTCTGACTCTCAGCTGACTGATGACGA 1963
 QY 1683 CTTGTCAACCTGACAGACTTCCACACCCAGTTCCTATGATACAGTTATCAAAAGATATT 1742
 DB 1964 CTTGTCAACCTGACAGACTTCCACACCCAGTTCCTATGATACAGTTATCAAAAGATATT 2023
 QY 1743 ATCAATTCCTCCCTGATGAGAGATTTGTGGAATTTTCAAGGCTTCAAGTTATTCCA 1802
 DB 2024 ATCAATTCCTCCCTGATGAGAGATTTGTGGAATTTTCAAGGCTTCAAGTTATTCCA 2083
 QY 1803 CTATTTCAAAAGAAATATGATTTAGTTTGGGGGGTGTGAGCTATGTTCAAAAGATGAG 1862
 DB 2084 CTATTTCAAAAGAAATATGATTTAGTTTGGGGGGTGTGAGCTATGTTCAAAAGATGAG 2143
 QY 1863 AACAGCTTGTGCTACATCTTCCACTTCTCAGCTCTCTCAGCTGTGTACTGCTTG 1922
 DB 2144 AACAGCTTGTGCTACATCTTCCACTTCTCAGCTCTCTCAGCTGTGTACTGCTTG 2203
 QY 1923 CAAAGACCCGGGAGCTGGCGGGGAAACCTTGGAGATGCTAGTTTGGCTTTTGGCTGAC 1982
 DB 2204 CAAAGACCCGGGAGCTGGCGGGGAAACCTTGGAGATGCTAGTTTGGCTTTTGGCTGAC 2263
 QY 1983 GAGAGGCTATGTTAAACAAACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2042
 DB 2264 GAGAGGCTATGTTAAACAAACACAGCAGAGATCGAAGGAGTTTGTAGAGATGTGTTCAA 2323

```

Db 764 CACGATCTCCAGGCGCTTATATGCGCGTTTGATACGATGGATGAAGCAACCATG 823
Qy 543 TGTGGATGTGGACGAGTGTGCAACAGATTCCACAGTGCACCCCAACGATCTGCAT 602
Db 824 TGTGGATGTGGACGAGTGTGCAACAGATTCCACAGTGCACCCCAACGATCTGCAT 883
Qy 603 CAATCTGAGGCGGGGTACACCTGCTCTGACCGAGGATATTGGTTTGGAAAGGCA 662
Db 884 CAATCTGAGGCGGGGTACACCTGCTCTGACCGAGGATATTGGTTTGGAAAGGCA 943
Qy 663 GTGCTTAGACATGTGATGTGCGCTATGTTACTGCGACAGACTCTGTGCGAATGTTCC 722
Db 944 GTGCTTAGACATGTGATGTGCGCTATGTTACTGCGACAGACTCTGTGCGAATGTTCC 1003
Qy 723 TGGATCTTATTTTGTACATGCAACCGCTGTTTACCCTTAATGAGATGGAAGTCTTG 782
Db 1004 TGGATCTTATTTTGTACATGCAACCGCTGTTTACCCTTAATGAGATGGAAGTCTTG 1063
Qy 783 CCAAGATGGAACGAGTGTGCCACGAGAACCCCTGGGTGCAAACTGGCGTCAACACTA 842
Db 1064 CCAAGATGGAACGAGTGTGCCACGAGAACCCCTGGGTGCAAACTGGCGTCAACACTA 1123
Qy 843 CGGCTCTTTCATCTGCGCGCTGTGACCCAGAGATATGAATGAGGAAAGTGGCTTCATG 902
Db 1124 CGGCTCTTTCATCTGCGCGCTGTGACCCAGAGATATGAATGAGGAAAGTGGCTTCATG 1183
Qy 903 CAGTATATGAGCAGAGTGTGCTCTGAGTTCCTTGGCAACATGAGTGTGAACCA 962
Db 1184 CAGTATATGAGCAGAGTGTGCTCTGAGTTCCTTGGCAACATGAGTGTGAACCA 1243
Qy 963 GCCCGGCAATCTTCTGCTGCGCGCTGACAGCTACATCCTGCTGATGACCAACGAG 1022
Db 1244 GCCCGGCAATCTTCTGCTGCGCGCTGACAGCTACATCCTGCTGATGACCAACGAG 1303
Qy 1023 CTGCCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCACACTGCAGACAGTG 1082
Db 1304 CTGCCAAGACATCAACGAATGTGAGACAGAAACACAGCTGCACACTGCAGACAGTG 1363
Qy 1083 CTACATATTACAAAGGGGCTCTCAAAATGACATCGGAGCCCATCGCTGTAGAGGCTTAACT 1142
Db 1364 CTACATATTACAAAGGGGCTCTCAAAATGACATCGGAGCCCATCGCTGTAGAGGCTTAACT 1423
Qy 1143 GAGGATCAGATGAATACCGCTGTATGTCCTGCTGAGAACCCCTGGCTGACAGACAGCC 1202
Db 1424 GAGGATCAGATGAATACCGCTGTATGTCCTGCTGAGAACCCCTGGCTGACAGACAGCC 1483
Qy 1203 CTTTACCATCTTGTACCGGAGACATGAGACGTGTGTCAGACGCTCCGTTCCCGCTGACAT 1262
Db 1484 CTTTACCATCTTGTACCGGAGACATGAGACGTGTGTCAGACGCTCCGTTCCCGCTGACAT 1543
Qy 1263 CTTCCAAATGCAAGCCACGACCCGCTAACCTGGGGCTTATTAATTTTCCAGATCAATC 1322
Db 1544 CTTCCAAATGCAAGCCACGACCCGCTAACCTGGGGCTTATTAATTTTCCAGATCAATC 1603
Qy 1323 TGGGATGAGGGGAGGAATTTTACATGGGGCAAAAGGGGCGCCATCAGTGCACCCCTGCT 1382
Db 1604 TGGGATGAGGGGAGGAATTTTACATGGGGCAAAAGGGGCGCCATCAGTGCACCCCTGCT 1663
Qy 1383 GATGACAGCCGCCATCAAAAGGGGCGCGGAAATCCAGCTGGACTTGGAAATGATCACTGT 1442
Db 1664 GATGACAGCCGCCATCAAAAGGGGCGCGGAAATCCAGCTGGACTTGGAAATGATCACTGT 1723
Qy 1443 CAACACTGTCAATCACTTCAAGAGGAGCTCCGATCCGAGCTGGGATATATGTGTCCGA 1502
Db 1724 CAACACTGTCAATCACTTCAAGAGGAGCTCCGATCCGAGCTGGGATATATGTGTCCGA 1783
Qy 1503 GTACCATCTGTGAGCTCGGGCTGAGAGCTCCGAGCGCTGCTCATTTGAGCAACGAAGGA 1562
Db 1784 GTACCATCTGTGAGCTCGGGCTGAGAGCTCCGAGCGCTGCTCATTTGAGCAACGAAGGA 1843
Qy 1563 CAGGAGAGAGAGGAATTAACAGAGAGAAATGAGAGCAGACAGAGCTTGGCATTTCTCG 1622

```

```

Db 1844 CAGGAGAGAGAGGAATTAACAGAGAGAAATGAGAGCAGACAGAGCTTGGCATTTCTCG 1903
Qy 1623 CTGACGTTTCCCGGAGAGTGCAGCCCGGACTCCCTGACTCTGACTGATATTGCGA 1682
Db 1904 CTGACGTTTCCCGGAGAGTGCAGCCCGGACTCCCTGACTCTGACTGATATTGCGA 1963
Qy 1683 CCTGTACCCCTGAGAGACTTGCACCCCGGAGTTCCTATGATACAGTATCAAAAAAGTAT 1742
Db 1964 CCTGTACCCCTGAGAGACTTGCACCCCGGAGTTCCTATGATACAGTATCAAAAAAGTAT 2023
Qy 1743 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTTCAAGGCTTCAGTTATTTCCA 1802
Db 2024 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTTCAAGGCTTCAGTTATTTCCA 2083
Qy 1803 CTATTTTCAAGAAATAGATAGTGTGTGGGGGCTGAGTCTATGTCAAGACTGTG 1862
Db 2084 CTATTTTCAAGAAATAGATAGTGTGTGGGGGCTGAGTCTATGTCAAGACTGTG 2143
Qy 1863 AACAGCTGCTGCTACCTTCTTCACTCTTCCACTCTCTCTCACTGTGTACTGCTTG 1922
Db 2144 AACAGCTGCTGCTACCTTCTTCACTCTTCCACTCTCTCTCACTGTGTACTGCTTG 2203
Qy 1923 CAAAGACCCGGAGCTGGCGGGAGACCCCTGGAGTACCTGATTTGCTTTTGGGACCA 1982
Db 2204 CAAAGACCCGGAGCTGGCGGGAGACCCCTGGAGTACCTGATTTGCTTTTGGGACCA 2263
Qy 1983 GAGAGGCTATGTAAACAAACCAACAGAGATCGAAGGGTTTATAGAGATGTGTCAA 2042
Db 2264 GAGAGGCTATGTAAACAAACCAACAGAGATCGAAGGGTTTATAGAGATGTGTCAA 2323
Qy 2043 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTTCAAGTTGCTTAAATTTGTAAAC 2102
Db 2324 AACCATGCTGTGATTTTCAACCATTAAGAAAGTTTCAAGTTGCTTAAATTTGTAAAC 2383
Qy 2103 GGTATATCTGTCTGTATTTTGTAGTATTTTAAATAATGCTGATTAATTCCTTC 2162
Db 2384 GGTATATCTGTCTGTATTTTGTAGTATTTTAAATAATGCTGATTAATTCCTTC 2443
Qy 2163 GAAAGGCTTCAGACACATGCTATGTTCTGCTCCCAACCCAGTCTCTCCATTTT 2222
Db 2444 GAAAGGCTTCAGACACATGCTATGTTCTGCTCCCAACCCAGTCTCTCCATTTT 2503
Qy 2223 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTGTCTTCTTGAATTTTACCATTT 2282
Db 2504 AGCCAGTGTCTTCTTGAAGACCCCTTAATCTTGTCTTCTTGAATTTTACCATTT 2563
Qy 2283 GCATTGGAATGACAGAGTCTCCCAACGATTAATTTTGAAGAGA 2328
Db 2564 GCATTGGAATGACAGAGTCTCCCAACGATTAATTTTGAAGAGA 2609

```

```

RESULT 6
US-10-176-758-407
; Sequence 407, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C104
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file Wrapper or Palm

```



```
DB 1424 GAGGATGAGTAAACCGCTGATGTGCTGCTGAGAACCTGGCTGAGAGACGACC 1483
QY 1203 CTTTACCATCTTGACCGGGGACATGAGAGTGTGCTGAGAGCGCTCCGTTCCGCTGACAT 1262
DB 1484 CTTTACCATCTTGACCGGGGACATGAGAGTGTGCTGAGAGCGCTCCGTTCCGCTGACAT 1543
QY 1263 CTTTCAAAATGCAAGCCAGACCGCTGAGAGCGCTGAGAGCGCTGAGAGCGCTGAGAGCGCT 1322
DB 1544 CTTTCAAAATGCAAGCCAGACCGCTGAGAGCGCTGAGAGCGCTGAGAGCGCTGAGAGCGCT 1603
QY 1323 TGGGAATGAGGGGACGAGAAATTTACATGCGGCAAAAGGGGCCCAATGTCACACCTGTG 1382
DB 1604 TGGGAATGAGGGGACGAGAAATTTACATGCGGCAAAAGGGGCCCAATGTCACACCTGTG 1663
QY 1383 GATGACACGGCCCATCAAAAGGGGCCCGGAAATCCAGCTGAGCTTGGAATGATCCTGT 1442
DB 1664 GATGACACGGCCCATCAAAAGGGGCCCGGAAATCCAGCTGAGCTTGGAATGATCCTGT 1723
QY 1443 CAACACTGTCATCAACTTTCAGAGGACAGCTCCGTGATCCGAGCTGCGGATATATGTGCGCA 1502
DB 1724 CAACACTGTCATCAACTTTCAGAGGACAGCTCCGTGATCCGAGCTGCGGATATATGTGCGCA 1783
QY 1503 GTTACCATCTTGAGCTTGGGCTGGAGCTTCGACGCTGCTGCTGATGGCAACCAAGGA 1562
DB 1784 GTTACCATCTTGAGCTTGGGCTGGAGCTTCGACGCTGCTGCTGATGGCAACCAAGGA 1843
QY 1563 CAGGAGAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGAGGTTAGGATTTCCG 1622
DB 1844 CAGGAGAGAGAGAAATTAACAGAGAGATGAGAGCGACACAGAGGTTAGGATTTCCG 1903
QY 1623 CTGAACTTTCCCGGAGAGTACAGCCCGACTCTCTGACTCTCACCTGACTACTATTGCGA 1682
DB 1904 CTGAACTTTCCCGGAGAGTACAGCCCGACTCTCTGACTCTCACCTGACTACTATTGCGA 1963
QY 1683 CCGTGCACCTGCGAGGACTTGCACCCGACTTCTTATGATACAGTATCAAAAAGTAT 1742
DB 1964 CCGTGCACCTGCGAGGACTTGCACCCGACTTCTTATGATACAGTATCAAAAAGTAT 2023
QY 1743 ATCATGCTCCCGGATGAGAAATGTTGTTGATTTTCAAGGCGCTGCTGATTTTCA 1802
DB 2024 ATCATGCTCCCGGATGAGAAATGTTGTTGATTTTCAAGGCGCTGCTGATTTTCA 2083
QY 1803 CTATTTTCAAGAAATAGATAGTTGCGGGGCTGAGTCTAATGTTCAAAAGACTGTG 1862
DB 2084 CTATTTTCAAGAAATAGATAGTTGCGGGGCTGAGTCTAATGTTCAAAAGACTGTG 2143
QY 1863 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1922
DB 2144 AAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2203
QY 1923 CAAGAAGCCCGGAGCTGGGCGGAAACCTGGGAGTGTGTTGCTTTTCCGTTGCGA 1982
DB 2204 CAAGAAGCCCGGAGCTGGGCGGAAACCTGGGAGTGTGTTGCTTTTCCGTTGCGA 2263
QY 1983 GAGAGGCTATGTTAAACAAACACAGAGATCGAAGGGTTTGAAGATGTTTCA 2042
DB 2264 GAGAGGCTATGTTAAACAAACACAGAGATCGAAGGGTTTGAAGATGTTTCA 2323
QY 2043 AACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2102
DB 2324 AACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2383
QY 2103 GGTTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2162
DB 2384 GGTTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2443
QY 2163 GAAAGGCTTACAGACATGCTATGTTGCTGCTTCCAAACCCAGTCTCTCTCAATTT 2222
DB 2444 GAAAGGCTTACAGACATGCTATGTTGCTGCTTCCAAACCCAGTCTCTCTCAATTT 2503
QY 2223 AGCCCATGTTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCAAT 2282
DB 2504 AGCCCATGTTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCAAT 2563
QY 2283 GGATTGGAATGACAGAGCTTCCAAACCTGATTAATTAATTTGAAGAGA 2328
DB 2564 GGATTGGAATGACAGAGCTTCCAAACCTGATTAATTAATTTGAAGAGA 2609

RESULT 5
US-10-174-590-407
: Sequence 407, Application US/10174590
: Publication No. US2003008352A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Chen, Jian
: APPLICANT: Desnoyers, Luc
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Pan, James
: APPLICANT: Smith, Victoria
: APPLICANT: Watanabe, Collin K.
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3430R1CA2
: CURRENT APPLICATION NUMBER: US/10/174,590
: PRIORITY FILING DATE: 2002-06-18
: PRIORITY APPLICATION removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 612
: SEQ ID NO 407
: LENGTH: 2609
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-174-590-407

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGCGGCTGCTCCCGTCTCTCTCTGACGACTGCTGCGCCCTCTGGAATTAACACCC 62
DB 284 CCGCGGCTGCTCCCGTCTCTCTCTGACGACTGCTGCGCCCTCTGGAATTAACACCC 343
QY 63 GCGAGCCCGGAGGCGCCAGAGAGGCGAGCGTCCGCGGAGTCTCCGCGGAGTCCCGCG 122
DB 344 GCGAGCCCGGAGGCGCCAGAGAGGCGAGCGTCCGCGGAGTCTCCGCGGAGTCCCGCG 403
QY 123 CGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182
DB 404 CGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 463
QY 183 AAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
DB 464 AAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523
QY 243 ACACTGACGAAATGCTTGAACCTGATGCGCAGTCCAGAGACAGTGTGATATGATGA 302
DB 524 ACACTGACGAAATGCTTGAACCTGATGCGCAGTCCAGAGACAGTGTGATATGATGA 583
QY 303 ATGCGGAAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGCGG 362
DB 584 ATGCGGAAACATCCCGGAGGCTGCGGAGAGACATGATGTGTTAAACCAAAATGCGG 643
QY 363 GTATTTAATGCTTCCCGGAGCAAAACCTGATGATGAGAGGCGCTTACTGCAACCCCTACT 422
DB 644 GTATTTAATGCTTCCCGGAGCAAAACCTGATGATGAGAGGCGCTTACTGCAACCCCTACT 703
QY 423 GACCCCTTACTCAGATCGTACACAGACAGCTGCCCCACACATCTCAGTCCAAACTATCC 482
DB 704 GACCCCTTACTCAGATCGTACACAGACAGCTGCCCCACACATCTCAGTCCAAACTATCC 763
QY 483 CACGATCTCCAGGCGCTTATATATGCGCGCTTGGATACAGATGATGAAGCAACCAATG 542
```

PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 99.8%; Score 2322.8; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCGGCGCTCTCCCGTGTCTCTCCACGACGCTGCGTCCCGCTCTGGAATAAACACCC 62
DB 284 CCGGCGCTCTCCCGTGTCTCTCTCCACGACGCTGCGTCCCGCTCTGGAATAAACACCC 343

QY 63 GCGAGCCCGGAGGCGCCAGAGAGGCGGACGTCGCGGAGCTCTCCGGGGGTCGCCCGC 122
DB 344 GCGAGCCCGGAGGCGCCAGAGAGGCGGACGTCGCGGAGCTCTCCGGGGGTCGCCCGC 403
QY 123 CGAGCTTCTCTCTCGCTTCTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
DB 404 CGAGCTTCTCTCTCGCTTCTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 463
QY 183 AAGGATACCTCTGATACCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
DB 464 AAGGATACCTCTGATACCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
QY 243 ACAGTCAAGATGCTTACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 302
DB 524 ACAGTCAAGATGCTTACCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 583
QY 303 ATGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGCTGATGCTGATGCTGATGCTG 362
DB 584 ATGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGCTGATGCTGATGCTGATGCTG 643
QY 363 GTATTATGATTCCTCCCGGACAAACCTGCTGATGCTGATGCTGATGCTGATGCTGAT 422
DB 644 GTATTATGATTCCTCCCGGACAAACCTGCTGATGCTGATGCTGATGCTGATGCTGAT 703
QY 423 GACCCCTACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482
DB 704 GACCCCTACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763
QY 483 CACGATCTCAGGCTCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
DB 764 CACGATCTCAGGCTCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
QY 543 TGTGATGTGAGAGAGTGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
DB 824 TGTGATGTGAGAGAGTGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 883
QY 603 CAATCTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
DB 884 CAATCTGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
QY 663 GTGCTTAGCATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
DB 944 GTGCTTAGCATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
QY 723 TCGATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
DB 1004 TCGATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
QY 783 CCAAGATGTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
DB 1064 CCAAGATGTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
QY 843 CGGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
DB 1124 CGGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183
QY 903 CAGTGTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
DB 1184 CAGTGTATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
QY 963 GCCCGGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
DB 1244 GCCCGGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303
QY 1023 CTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
DB 1304 CTGCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1363
QY 1083 CTGACATTTACAGAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
DB 1364 CTGACATTTACAGAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1423
QY 1143 GAGGATCACTGATTAACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202

Db 2436 AGCCAGGTTTCTTTAGAGCCCTTAATCTTCTTTAGATTTTACCAAT 2495
Oy 2283 GGATTGGAATGACAGAGTCTCCAACTGATTAATATTTGAGAGA 2328
Db 2496 GGATTGGAATGACAGAGTCTCCAACTGATTAATATTTGAGAGA 2541

RESULT 4

US-10-066-500-14
Sequence 14, Application US/10066500
Patent No. US20020177165A1
GENERAL INFORMATION:
APPLICANT: Avi J. Ashkenazi
APPLICANT: Kevin P. Baker
APPLICANT: David A. Botstein
APPLICANT: Luc Desnoyers
APPLICANT: Dan L. Eaton
APPLICANT: Napoleone Ferrara
APPLICANT: Sherman Fong
APPLICANT: Wei-Olang Gao
APPLICANT: Hanspeter Gerber
APPLICANT: Mary E. Gerltsen
APPLICANT: Audrey Goddard
APPLICANT: Paul J. Godowski
APPLICANT: Austin L. Gurney
APPLICANT: Ivar J. Kljavin
APPLICANT: Jennie P. Mather
APPLICANT: Mary A. Napier
APPLICANT: James Pan
APPLICANT: Nicholas F. Paoni
APPLICANT: Margaret Ann Roy
APPLICANT: Timothy A. Stewart
APPLICANT: Daniel Tumas
APPLICANT: Colin K. Watanabe
APPLICANT: P. Mickey Williams
APPLICANT: William T. Wood
APPLICANT: Zemin Zang
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P313081C7
CURRENT APPLICATION NUMBER: US/10/066,500
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 10/002,796
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063664
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/066964
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137

QY	63	CCGACACCCCGAGAGGCCGACGTCGCGAGCTCTCGGGGGTCCGCCG	122
Db	276	GCAGACCCCGAGAGGCCGAGAGGCCACACGCGCCGAGCTCTCCGGGGGTCCGCCG	335
QY	123	CGAGCTTCTTCCTTCGTCATGCTCCCTCGGGCGCTTGACATGCGCAGAAATAA	182
Db	336	CGAGCTTCTTCCTTCGTCATGCTCCCTCGGGCGCTTGACATGCGCAGAAATAA	395
QY	183	AAGATTACTCACTGTATTACCATTTCTGGCTCTCTCTTCCAAAGCCCTGGGAATGACAGGC	242
Db	396	AAGATTACTCACTGTATTACCATTTCTGGCTCTCTCTTCCAAAGCCCTGGGAATGACAGGC	455
QY	243	ACAGGACGAGATGGCTTTGACCTGGATCGGCAGTCAAGAGATGTTATATATGATGA	302
Db	456	ACAGGACGAGATGGCTTTGACCTGGATCGGCAGTCAAGAGATGTTATATATGATGA	515
QY	303	ATGCCGAACATCCCCGAGGCGCTCGCAGAGAGACATGATGTGTCTTAAACCAAAATGGCG	362
Db	516	ATGCCGAACATCCCCGAGGCGCTCGCAGAGAGATGATGTGTCTTAAACCAAAATGGCG	575
QY	353	GTATTATTGCATTCGCCGAGACAAACCTGTGTATGAGAGGCCCTACTGAAACCCCTACTC	422
Db	576	GTATTATTGCATTCGCCGAGACAAACCTGTGTATGAGAGGCCCTACTGAAACCCCTACTC	635
QY	423	GACCCCTACACAGGTCCGTACCCGACGAGCTGCCCCACACCTCAGCTCCAAATATACC	482
Db	636	GACCCCTACACAGGTCCGTACCCGACGAGCTGCCCCACACCTCAGCTCCAAATATACC	695
QY	483	CACGATCTCCAGGCGCTTATATATGCGGCTTTGGATACCAAGATGATGTAAGCAACCAATG	542
Db	696	CACGATCTCCAGGCGCTTATATATGCGGCTTTGGATACCAAGATGATGTAAGCAACCAATG	755
QY	543	TGTGGATGTGGACAGGTGTGCAACGATTTCCACACAGTGCACACCCACCCAGATTCGAT	602
Db	756	TGTGGATGTGGACAGGTGTGCAACGATTTCCACACAGTGCACACCCACCCAGATTCGAT	815
QY	603	CAATACTGAAGCGCGGTACACCTGCTCTCTGCACGACGAGATTTGGCTTCGGAAGGCA	662
Db	816	CAATACTGAAGCGCGGTACACCTGCTCTCTGCACGAGGATTTGGCTTCGGAAGGCA	875
QY	663	GTGCTTAGACATTGATGAATGCTCGTATGTTACTGCCAGCAGCTCTGTGCGAATGTTCC	722
Db	876	GTGCTTAGACATTGATGAATGCTCGTATGTTACTGCCAGCAGCTCTGTGCGAATGTTCC	935
QY	723	TGATATCCATTTCTGTATCATCAACCCGGTTTTACCTCAATGAGATGGAAGTCTGTG	782
Db	936	TGATATCCATTTCTGTATCATCAACCCGGTTTTACCTCAATGAGATGGAAGTCTGTG	995
QY	783	CCAAGATGTGAACGAGTGTGCCACCGAAMCCCTGCGTGCACAACTTCGTCAACACTTA	842
Db	996	CCAAGATGTGAACGAGTGTGCCACCGAAMCCCTGCGTGCACAACTTCGTCAACACTTA	1055
QY	843	CGGCTCTTTCATCTGCGCGCTGTGACCCAGGATATGAACTTAGAGAAATGCGCTTCATTTG	902
Db	1056	CGGCTCTTTCATCTGCGCGCTGTGACCCAGGATATGAACTTAGAGAAATGCGCTTCATTTG	1115
QY	903	CAGTATATGACAGATGACACTTCTGTCGATTCCTCTGCCAACAATGATGTGTGAACCA	962
Db	1116	CAGTATATGACAGATGACACTTCTGTCGATTCCTCTGCCAACAATGATGTGTGAACCA	1175
QY	963	GCCCGGCACATACCTTCTGCTCTGCGCTCCAGGCTACATCCTGTGTGATGACAAACCGAAG	1022
Db	1176	GCCCGGCACATACCTTCTGCTCTGCGCTCCAGGCTACATCCTGTGTGATGACAAACCGAAG	1235
QY	1023	CTGCCAAGACATCAACGAAATGTGACACAGAACCAACAGTGCACACCTGACGACAGCTG	1082
Db	1236	CTGCCAAGACATCAACGAAATGTGACACAGAACCAACAGTGCACACCTGACGACAGCTG	1295
QY	1083	CTACAATTTACAAGGGGCTTCAAAATGATGAGACCCCATCCGCTGTGAGGAGCCCTTATCT	1142
Db	1296	CTACAATTTACAAGGGGCTTCAAAATGATGAGACCCCATCCGCTGTGAGGAGCCCTTATCT	1355

QY	1143	CAGGATTCAGTATTAACGGCTGATGTGTCGCTGAGAACCCCTGGCTGCACAGACACC	1202
Db	1356	GAGGATTCAGTATTAACCCCTGTATGTGTCGCTGAGAACCCCTGGCTGCACAGACACC	1415
QY	1203	CTTTACCATCTTGTACCCGGACATGAGCGTGGTGTCCAGAGCGTCCGTTCCGGTGCAT	1262
Db	1416	CTTTACCATCTTGTACCCGGACATGAGCGTGGTGTCCAGAGCGTCCGTTCCGGTGCAT	1475
QY	1253	CTTCCAAATGCAAGCCAGACCCGCTACCCTGGGGCCCTATACATTTTCCAGATCAATC	1322
Db	1476	CTTCCAAATGCAAGCCAGACCCGCTACCCTGGGGCCCTATACATTTTCCAGATCAATC	1535
QY	1323	TGGGAATGAGGGACAGAGATTTTATCATCGGCCAAACGGGCCCATCACTAGTGCACCTGGT	1382
Db	1536	TGGGAATGAGGGACAGAGATTTTATCATCGGCCAAACGGGCCCATCACTAGTGCACCTGGT	1595
QY	1383	GATGACACGCCCCATCAAAAGGGCCCCGGGAAATCAGTGGACCTTGGAAATGATCAGCT	1442
Db	1596	GATGACACGCCCCATCAAAAGGGCCCCGGGAAATCAGTGGACCTTGGAAATGATCAGCT	1655
QY	1443	CAACACTGTCAATCACTTCAAGAGCAGCTCCGTATCCGACCTGGCGATATATGTGCGCA	1502
Db	1656	CAACACTGTCAATCACTTCAAGAGCAGCTCCGTATCCGACCTGGCGATATATGTGCGCA	1715
QY	1503	GTACCCATTTGAGCCTTCGGGCTGGAGCCTCCGACGCTCCCTCTCATTTGGCACCAGGGA	1562
Db	1716	GTACCCATTTGAGCCTTCGGGCTGGAGCCTCCGACGCTCCCTCTCATTTGGCACCAGGGA	1775
QY	1563	CAGGGAAGAAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTG	1622
Db	1776	CAGGGAAGAAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTG	1835
QY	1623	CTGAACGTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCCTCACTCACTGTACTATTGAGA	1682
Db	1836	CTGAACGTTTCCCGAAGAGTCAAGCCCGACTTCTGACTCCTCACTCACTGTACTATTGAGA	1895
QY	1683	CCTGCACCCCTGCAGAGACTTGCACACCCCACTTCTATGATACGATTAATCAAAAGTAT	1742
Db	1896	CCTGCACCCCTGCAGAGACTTGCACACCCCACTTCTATGATACGATTAATCAAAAGTAT	1955
QY	1743	ATCATTTGCTCCCGATAGAAAGATTTGTGTGAATTTTCAAGGCGCTCAGTTTATTTCCA	1802
Db	1956	ATCATTTGCTCCCGATAGAAAGATTTGTGTGAATTTTCAAGGCGCTCAGTTTATTTCCA	2015
QY	1803	CTATTTTAAAGAAATATGATTTAGTTTGGCGGGCTGTGAGTCTATGTTCAAGACGTGTG	1862
Db	2016	CTATTTTAAAGAAATATGATTTAGTTTGGCGGGCTGTGAGTCTATGTTCAAGACGTGTG	2075
QY	1863	AACAGCTTGCCTGACTCTTCACGCTCTTCACACCTCTCTCACTGCTGTTACTGCTTGTG	1922
Db	2076	AACAGCTTGCCTGACTCTTCACGCTCTTCACACCTCTCTCACTGCTGTTACTGCTTGTG	2135
QY	1923	CAAAAGCCCGGAGCTGGCGGGGAACCTGGGAATAGCTAGTTTGCCTTGTGCGTACACA	1982
Db	2136	CAAAAGCCCGGAGCTGGCGGGGAACCTGGGAATAGCTAGTTTGCCTTGTGCGTACACA	2195
QY	1983	GAGAAAGCTATGTAACAAACCAACACAGACAGATCGAAGGCTTTTAAAGAAATGTGTTCAA	2042
Db	2196	GAGAAAGCTATGTAACAAACCAACACAGACAGATCGAAGGCTTTTAAAGAAATGTGTTCAA	2255
QY	2043	AACCATGCGCTGTATTTTCAACCCATAAAGAAAGTTTCAAGTGTGCTTAAATTTGTATAC	2102
Db	2256	AACCATGCGCTGTATTTTCAACCCATAAAGAAAGTTTCAAGTGTGCTTAAATTTGTATAC	2315
QY	2103	GGTTTAATTCGTCTTGTTCATTTTGAATTTTTTAAAAAATATGCTGATGATTTCTCTC	2162
Db	2316	GGTTTAATTCGTCTTGTTCATTTTGAATTTTTTAAAAAATATGCTGATGATTTCTCTC	2375
QY	2163	GAAAGGCGCTACAGACATGCTATGTCGTGCTCCCAAAACCGACGTCCCTCCATTTT	2222
Db	2376	GAAAGGCGCTACAGACATGCTATGTCGTGCTCCCAAAACCGACGTCCCTCCATTTT	2435
QY	2223	AGCCAGTGTTTCTTTGAGAGACCCCTTAATCTTGCTTCTTTTGAATTTTTTACCAAT	2282

QY 1023 CTGCCAAGACATCAACGATGTGACACAGGAAACACAGTGCACACTGCACAGACGTG 1082
 DB 1040 CTGCCAAGACATCAACGATGTGACACAGGAAACACAGTGCACACTGCACAGACGTG 1099
 QY 1083 CTACAAATTTACAAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGACCTTATCT 1142
 DB 1100 CTACAAATTTACAAAGGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGACCTTATCT 1159
 QY 1143 GAGGATAGATGATACCGCTGTATGTCTCTCTGAGAAACCTGGCTGAGAGACAGCC 1202
 DB 1160 GAGGATAGATGATACCGCTGTATGTCTCTCTGAGAAACCTGGCTGAGAGACAGCC 1219
 QY 1203 CTTTACCATCTTGTACCGGGGACATGAGACGTGTCTCAGAGACCTCCGCTCCGCTGACAT 1262
 DB 1220 CTTTACCATCTTGTACCGGGGACATGAGACGTGTCTCAGAGACCTCCGCTCCGCTGACAT 1279
 QY 1263 CTTTCAAAATGACAGCCACGACCCCGCTACCTGGGGGCTTATTAATTTTCCAGATCAATTC 1322
 DB 1280 CTTTCAAAATGACAGCCACGACCCCGCTACCTGGGGGCTTATTAATTTTCCAGATCAATTC 1339
 QY 1323 TGGGAATGAGGGGACAGATTTTACATGCGGAAACGGGGCCCATCAGTGCACACCTGGT 1382
 DB 1340 TGGGAATGAGGGGACAGATTTTACATGCGGAAACGGGGCCCATCAGTGCACACCTGGT 1399
 QY 1383 GATGACACGCCCCCATCAAGGGGCCCCGGGAAATCCAGCTGAGACTTGGAAATGATCACTGT 1442
 DB 1400 GATGACACGCCCCCATCAAGGGGCCCCGGGAAATCCAGCTGAGACTTGGAAATGATCACTGT 1459
 QY 1443 CAACACTGTATCAACTTCAAGAGCAGCTCCGTATCCGACTGCGGATATATGTGTGCA 1502
 DB 1460 CAACACTGTATCAACTTCAAGAGCAGCTCCGTATCCGACTGCGGATATATGTGTGCA 1519
 QY 1503 GTACCCATCTGAGCTCGGGCTGGAGCTCCGAGCTGCGCTCTGATTTGGACCAAGGGA 1562
 DB 1520 GTACCCATCTGAGCTCGGGCTGGAGCTCCGAGCTGCGCTCTGATTTGGACCAAGGGA 1579
 QY 1563 CAGGAGAAGAGAGAAATACAGAGAGATGAGAGACGACACAGAGCTTATGAGCTTCTG 1622
 DB 1580 CAGGAGAAGAGAGAAATACAGAGAGATGAGAGACGACACAGAGCTTATGAGCTTCTG 1639
 QY 1623 CTGACACTTCCCGGAGAGCTGACCGGCTCTGACTCTCAGCTGACTGACTGACTGACTG 1682
 DB 1640 CTGACACTTCCCGGAGAGCTGACCGGCTCTGACTCTCAGCTGACTGACTGACTGACTG 1699
 QY 1683 CCGTGCACCCGAGGAGCTTCCACCCCGAGTCTCTATGATACAGTATCAAAAGTAT 1742
 DB 1700 CCGTGCACCCGAGGAGCTTCCACCCCGAGTCTCTATGATACAGTATCAAAAGTAT 1759
 QY 1743 ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCAAGTATTTTCA 1802
 DB 1760 ATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCAAGTATTTTCA 1819
 QY 1803 CTATTTTCAAGAAATAGATAGTGTGCGGGGCTGTGATGTATGTCAAAAGACTGTG 1862
 DB 1820 CTATTTTCAAGAAATAGATAGTGTGCGGGGCTGTGATGTATGTCAAAAGACTGTG 1879
 QY 1863 AACAGCTGTGCTGACTTCTCAGCTCTCTCAGCTCTCTCAGCTGTGTTACTGTTG 1922
 DB 1880 AACAGCTGTGCTGACTTCTCAGCTCTCTCAGCTCTCTCAGCTGTGTTACTGTTG 1939
 QY 1923 CAAGAGCCCGGAGCTGCGGGGAAACCTGGGAGTATGATTTGCTTTTTCGTTAC 1982
 DB 1940 CAAGAGCCCGGAGCTGCGGGGAAACCTGGGAGTATGATTTGCTTTTTCGTTAC 1999
 QY 1983 GAGAGGCTATGTAAACAACACAGAGATGAGAGGTTTATGAGATGTGTTCA 2042
 DB 2000 GAGAGGCTATGTAAACAACACAGAGATGAGAGGTTTATGAGATGTGTTCA 2059
 QY 2043 AACCATGCTGTATTTTCAACATTAAGAGTTCAGTGTCTTAAATTTGTTAAC 2102
 DB 2060 AACCATGCTGTATTTTCAACATTAAGAGTTCAGTGTCTTAAATTTGTTAAC 2119

QY 2103 GGTAAATCTGCTGCTTGTCAATTTTGAATTTTAAAAAATATGCTAGATTTCTTC 2162
 DB 2120 GGTAAATCTGCTGCTTGTCAATTTTGAATTTTAAAAAATATGCTAGATTTCTTC 2179
 QY 2163 GAAAGGCTTACAGACATGCTATGTTCTGCTTCCAAACCAAGCTCTCTCATTTT 2222
 DB 2180 GAAAGGCTTACAGACATGCTATGTTCTGCTTCCAAACCAAGCTCTCTCATTTT 2239
 QY 2223 AGCCAGCTGTTTCTTTGAGAGACCCCTTAATCTGCTTCTTAAATTTTACCAAT 2282
 DB 2240 AGCCAGCTGTTTCTTTGAGAGACCCCTTAATCTGCTTCTTAAATTTTACCAAT 2299
 QY 2283 GGAATGGAATGACAGAGCTTCCAAACTGATTAATTAATTTTGAAGAGA 2328
 DB 2300 GGAATGGAATGACAGAGCTTCCAAACTGATTAATTAATTTTGAAGAGA 2345

RESULT 3
 US-09-836-561-2
 : Sequence 2, Application US/09836561
 : Patent No. US20020038006A1
 :
 : GENERAL INFORMATION:
 : APPLICANT: Handman, Olga
 : Corley, Neil C.
 : Guegler, Karl J.
 : TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 94304
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/836,561
 : FILING DATE: 16-Apr-2001
 : CLASSIFICATION: <Unknown>
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 09/212,168
 : FILING DATE: <Unknown>
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PF-0333 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-855-0555
 : TELEFAX: 415-845-4166
 : TELEX: <Unknown>
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 2550 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: CORNNOT01
 : CLONE: 45517
 : SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 : US-09-836-561-2

Query Match 99.8%; Score 2324.4; DB 10; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCGGGCCCTCTCCCGGCTGCTCTCCAGCACTGCTGCGCCCTCTGGAATTAACACCC 62
 DB 216 CCGGGCCCTCTCCCGGCTGCTCTCCAGCACTGCTGCGCCCTCTGGAATTAACACCC 275

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 11:41:03 ; Search time 379.284 Seconds

(Without alignments)
\$581.191 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328
Sequence: 1 gaccgcgcgcctccccgcg.....tgacttaatatitgagaga 2328

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 1085931 seqs, 780495707 residues

Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	99.9	2362	9	US-10-041-016-1
2	2326	99.9	2362	10	US-09-083-002-1
3	2324.4	99.8	2550	10	US-09-836-561-2
4	2322.8	99.8	2609	9	US-10-066-500-14
5	2322.8	99.8	2609	9	US-10-174-580-407
6	2322.8	99.8	2609	9	US-10-176-758-407
7	2322.8	99.8	2609	9	US-10-175-737-407
8	2322.8	99.8	2609	9	US-10-173-706-407
9	2322.8	99.8	2609	9	US-10-175-738-407
10	2322.8	99.8	2609	9	US-10-175-752-407
11	2322.8	99.8	2609	9	US-10-176-482-407
12	2322.8	99.8	2609	9	US-10-176-757-407
13	2322.8	99.8	2609	9	US-10-176-913-407
14	2322.8	99.8	2609	9	US-10-180-552-407
15	2322.8	99.8	2609	9	US-10-180-557-407
16	2322.8	99.8	2609	9	US-10-173-700-407
17	2322.8	99.8	2609	9	US-10-174-572-407
18	2322.8	99.8	2609	9	US-10-174-579-407
19	2322.8	99.8	2609	9	US-10-174-582-407

20	2322.8	99.8	2609	9	US-10-174-588-407
21	2322.8	99.8	2609	9	US-10-175-739-407
22	2322.8	99.8	2609	9	US-10-175-740-407
23	2322.8	99.8	2609	9	US-10-175-743-407
24	2322.8	99.8	2609	9	US-10-176-488-407
25	2322.8	99.8	2609	9	US-10-176-492-407
26	2322.8	99.8	2609	9	US-10-176-747-407
27	2322.8	99.8	2609	9	US-10-176-750-407
28	2322.8	99.8	2609	9	US-10-176-985-407
29	2322.8	99.8	2609	9	US-10-176-987-407
30	2322.8	99.8	2609	9	US-10-176-991-407
31	2322.8	99.8	2609	9	US-10-176-992-407
32	2322.8	99.8	2609	9	US-10-176-993-407
33	2322.8	99.8	2609	9	US-10-184-658-407
34	2322.8	99.8	2609	9	US-10-002-796-14
35	2322.8	99.8	2609	9	US-10-066-494-14
36	2322.8	99.8	2609	9	US-10-173-695-407
37	2322.8	99.8	2609	9	US-10-173-697-407
38	2322.8	99.8	2609	9	US-10-173-705-407
39	2322.8	99.8	2609	9	US-10-174-576-407
40	2322.8	99.8	2609	9	US-10-174-585-407
41	2322.8	99.8	2609	9	US-10-174-586-407
42	2322.8	99.8	2609	9	US-10-175-747-407
43	2322.8	99.8	2609	9	US-10-176-481-407
44	2322.8	99.8	2609	9	US-10-176-485-407
45	2322.8	99.8	2609	9	US-10-176-485-407

ALIGNMENTS

RESULT 1
US-10-041-016-1
Sequence 1, Application US/10041016
Patent No. US20020165151A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racie, Lisa A.
Lavallee, Edward R.
Merberg, David
Treacy, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasuku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851


```

Db 3181 CGAGATGTGACGAGTGTGCCAGCCGAGCCTTCGCCCCAGCGGCTTCCTCAACACG 3240
QY 841 TACGGCTCTTTCATCTGCGG---CTGTGACCCAGATATGAATCTGAGAGATGGCGTT 897
Db 3241 GAGGGCTCTTCACCTGCTCAGCCTGTGACAGCGGCTACTGGGTGAACGAGATGGCAGCT 3300
QY 898 CATTGCACTATATGACGAGTGCAGCTTCTCTGAGTTCCTCTGCCAATGATGTGTG 957
Db 3301 GCGCTGAAGACTGTGATGAATGTGCTTCCTGAGTGTGCCCCACAGAGCCTTCACAC 3360
QY 958 AACGAGCCGCGACATCTTCTGCTCTGCTCCAGGCTTACATCTGCTGTGATGACAC 1017
Db 3361 AATCTGTAGCTCTCTCTCTCTGCAAGAGACTGTACAGGGCTACCGGCCAACCCCTG 3420
QY 1018 CGAAGCTGCCAAGACATCAACGATGTGACACAGAACACACGTCGACCTGACAGAG 1077
Db 3421 GCGAACAGATGCGAAGATGTGATGATGTGTAAGTCCCAAGACGCTGCCGGGAGGC 3480
Y 1078 ACGTGTACATTTTACAGGGGGGCTTCAATGATCATGACCCCATCCGCTGTGAGAGCT 1137
Db 3481 GAATGCAAGAACACAGAAAGTCTTCAACATGCTCTGTACACAGGCTTCACAGCTGTC 3540
QY 1138 TATCTGAGGATCAGTATACCGCTGTATGTGTCTGCTGAGAACCTGG 1187
Db 3541 AATGGCACCATGTGTGAGGAGCGTGAATGATGTGTGGGGAAGACATTG 3590

RESULT 14
PCT-US95-02251-17
; Sequence 17, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UMIC009P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5502 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5502
; PCT-US95-02251-17
Query Match
2.3%; Score 52.4; DB 5; Length 5502;
Best Local Similarity 47.4%; Pred. No. 5e-05;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;

QY 724 GGAATCCATTTCTTGTACATGACAACTGTTTACCTCAATGATGAGATGCTTTC 783
Db 3121 GGTCTCTTAGATGCTCTCTGTGAGCGGGCTATGAGTCAACCCAGACAAAGAGGCTTC 3180
QY 784 CAAGATGTGACGAGTGTGCCACCGAAGCCCTG---CGTGAACCTCGTCAACACC 840
Db 3181 CGAGATGTGAGAGAGTGTCCACGCGAGCTCTGCCCCACAGGCGCTCTGCTCAACAG 3240
QY 841 TACGCTCTTTCATCTGCGG---CTGTGACCCAGATATGAACTTGAGAAATGGCGTT 897
Db 3241 GAGGCTCTTTCACCTGCTCAGAGCTGTCAAGAGGGGACTGGGTGAACGAAAGTGGCACT 3300
QY 898 CATTGCACTATATGAGAGAGTGCAGCTTCTCTGAGTTCCTCTGCCAATGATGTGTG 957
Db 3301 GCGCTGTGAAGACTGTGATGAATGTGCTTCCTGAGTGTGCCCCACAGCGCTGTGACAC 3360
QY 958 AACGAGCCGCGACATCTTCTGCTCCAGGCTTACATCTGCTGTGATGACAC 1017
Db 3361 AATCTGTAGCTCTCTCTCTGCAAGAGACTGTACAGGGCTACCGGCCAACCCCTG 3420
QY 1018 CGAAGCTGCCAAGACATCAAGATGTGACACAGAACACACGTCGACCTGACAGAG 1077
Db 3421 GCGAACAGATGCGAAGATGTGATGATGTGTAAGTCCCAAGACAGCTGCCGGGAGGC 3480
QY 1078 ACGTGTACATTTTACAAAGGGGCTTCAATGATGATGACCCCATCCGCTGTGAGAGCT 1137
Db 3481 GAATGCAAGAACACAGAAAGTCTTCAACATGCTCTGTACACAGGCTTCACAGCTGTC 3540
QY 1138 TATCTGAGGATCAGTATACCGCTGTATGTGTCTGCTGAGAACCTGG 1187
Db 3541 AATGGCACCATGTGTGAGGAGCGTGAATGATGTGTGGGGAAGACATTG 3590

RESULT 15
PCT-US95-02251-2
; Sequence 2, Application PC/TUS9502251
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02251
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; CLASSIFICATION:

```


DB 940 ACAAAATGA--ATGCCGGAGAGATGAATGTGTGAATATATCATGCGGCTTCGGTGT 996
QY 1111 ATCGACCCATCGCGTGTGAGAGCCCTATCTGAGAGTGAATACCGCTGTATGTGT 1170
DB 997 TATCCAGAAATCTTGTCTCAAGATCCCTACATCTTAACACCAAGAAACCGATGTGTTC 1056
QY 1171 CCTGCTGAGAACCCCTGGCTGACAGACCCCTTTACCATCTTGTACCGGACATGAGC 1230
DB 1057 CCAGTCTCAATTCCTCAATGTCCGAGAACTGCCCGCATATAGTCTACAAATACATGAGC 1116
QY 1231 GTGGTGTCAAGAGCGTCCGCTGCTCCGCTGACATCTTCCAAATGACAGCCGCGTAC 1290
DB 1117 ATCCGATCTGATAGGCTGTGTGCTGACATCTTCCAGATACAGGCGCAACAATATAT 1176
QY 1291 CCGGGGCGCTATATACATTTTCCAGATCAATCTGGAGATGAGGCGAGAAATTTACATG 1350
DB 1177 GCGAACACATCAATCTTGTGATTAATCTGGAATGAAATGAGAGTTCACCTA 1236
QY 1351 CGGCAAAAGGCGCCCATCATGATGCCACCCCTGATGACACGCCCATCAAGAGGCGCCG 1410
DB 1237 CGCAAAACAGTCTGTGAATGCAATGCTGTGCTGGAAGTCAATATCAAGAGACCAAGA 1296
QY 1411 GAATTCAGCTGACTTGAATATGATCATCTCAACACTGTCACTCACTTCAAGAGCAGC 1470
DB 1297 GAACATATGCTGAGCTGAGATGCTGACAGTACAGATATAGGAGACCTTCGCAACAAGC 1356
QY 1471 TCCGATATGCTGAGCTGAGATATATGTGCGAGTACCATCTT 1513
DB 1357 TCTGTGTTAAGATTGACAAATATATAGTGGGCGCATTTTCATTTT 1399

RESULT 8
US-08-897-443-2
Sequence 2, Application US/08897443
Patent No. 5981263
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: filed Herewith
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0348 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3373 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: Linear
IMMEDIATE SOURCE:
LIBRARY: UTRSN0702
CLONE: 681719
US-08-897-443-2

Query Match
Best Local Similarity 57.7%; Pred. No. 3.9e-21;
Matches 211; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

4.6%; Score 106.8; DB 2; Length 3373;
Best Local Similarity 57.7%; Pred. No. 3.9e-21;
Matches 211; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

QY 695 ACTGCCAGACGCTGTGCGAATGTTCTGATCTTCTATCTTGTACATGCAACCTGTG 754
DB 920 ACTGTGAGACGCTGTGTGATGATGCGGCGCTCCTGCTGCGCAGTGTACAGTGGCT 979
QY 755 TTACCCATGAGGAGGAGAGGCTTCCCAAGTGTACAGAGTGTGCGACCGAGAAC 814
DB 980 ACGCCCTGCTGAGAGAGGAGGAGTGTGTGCTGTGCTGACTACTGCTTACCTGCA 1039
QY 815 CC--TCCGTGCAAACTGCTGCAACACCTTACGCTTTCATCTGCGCTGTGACCGAG 871
DB 1040 ACGATCTGAACATGAGTGTGTAAATGCTGATGCTCTTACCTTGTGCGAGTGAAG 1099
QY 872 GATATGACTTGAGAGAGTGGCTTTCATTCAGTGTATGAGCAGTGTGACCTTCTG 931
DB 1100 GATTTGCTCTTAAACCCAGATTAATAAAGCTGCACAAATATGACTACTGTGCTCATCTA 1159
QY 932 AGTTCCTGCGCAACATGAGTGTGTGAGAACAGCCCGGCAATCTTGTGCTGCGCTC 991
DB 1160 ATACGGATGCTGACAGAGTGTGTATACAGATGATTTCTTCTGCGCAGTGTGCA 1219
QY 992 CAGGCTACATCTGCTGATGATGACCAACGAGCTGCGCAAGACATCAAGATGTGAGACA 1051
DB 1220 AAGCTTTACCTTAATTCACATTAAGAAACCTGCAAGAGATCAACTGTGTGACCTGA 1279
QY 1052 GGAACC 1057
DB 1280 ACAAAC 1285

RESULT 9
US-09-643-597-278
Sequence 278, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: US/09/643,597
CURRENT FILING DATE: 2000-08-21
NUMBER OF SEQ ID NOS: 369
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 278
LENGTH: 401
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(401)
OTHER INFORMATION: n - A,T,C or G
US-09-643-597-278

OY	2163	GAAGGCGCTTAGACACACATGCTATGTTCTGTCTTCCAAACCCAGTCTCCTCCATTTT	2222
Db	2376	GAAGGCGCTTAGACACACATGCTATGTTCTGTCTTCCAAACCCAGTCTCCTCCATTTT	2435
OY	2223	AGCCAGAGTTCTCTTTGAGGAGCCCTTAATCTGCTTTCITTTGGAATTTTACCCAATT	2282
Db	2436	AGCCAGAGTTCTCTTTGAGGAGCCCTTAATCTGCTTTCITTTGGAATTTTACCCAATT	2495
OY	2283	GGATTGGAATCAGAGGCTCTCCAACAGATTAATATTTTGAAGAGA	2328
Db	2496	GGATTGGAATCAGAGGCTCTCCAACAGATTAATATTTTGAAGAGA	2541

RESULT 2
US-09-212-168-2
; Sequence 2, Application US/09212168
; Patent No. 6303765

```

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517

```

Query Match	99.8%;	Score 2324.4;	DB 4;	Length 2550;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2325;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	Db	QY	Db
3	CCGGGCGCTCTCCCGGTCTCCTCTCAGACACACGCTCCCTCTGGAAATAAACACCC	62	
216	CCGGGCGCTCTCCCGGTCTCCTCTCAGACACGCTCCCTCTGGAAATAAACACCC	275	
63	GGGAGCCCCGAGGCGCCAGAGAGGCGACGTGCCGAGACTCTCCGGGGGTCCGCCCG	122	
276	GGGAGCCCCGAGGCGCCAGAGAGGCGACGTGCCGAGACTCTCCGGGGGTCCGCCCG	335	

QY	123	CGAGCTTCTCTCGCGCTTCGAGATCTCTCTCGCGCTTGGACATGCGAGGATATAA	182
Db	336	CGAGCTTCTTCTCGCCTTCGAGATCTCTCTCGCGCTTGGACATGCGAGGATATAA	395
QY	183	AAGGATACTCACTGTTACCATTCCTGCTCTCTCTCTCCAAAGCCCTGGGAATGCACAGCG	242
Db	396	AAGGATACTCACTGTTACCATTCCTGCTCTCTCTCTCCAAAGCCCTGGGAATGCACAGCG	455
QY	243	ACAGGACAGGATGGCTTGGACCTGATGCGCAGTACAGACAGTGTGTTAGATTGATGA	302
Db	456	ACAGGACAGGATGGCTTGGACCTGATGCGCAGTACAGACAGTGTGTTAGATTGATGA	515
QY	303	ATGCCGAACCATCCCCGAGGCGCTGCGGAGAGACATGATGTGTGTTAACCAAAATGGCGG	362
Db	516	ATGCCGAACCATCCCCGAGGCGCTGCGGAGAGACATGATGTGTGTTAACCAAAATGGCGG	575
QY	363	GTATTATTAGCATTTCCCGGACAAACCCCTGTATGAGAGGCGCCTACTGCAACCCCTACTC	422
Db	576	GTATTATTAGCATTTCCCGGACAAACCCCTGTATGAGAGGCGCCTACTGCAACCCCTACTC	635
QY	423	GACCCCTACTCAGGCTCGGTACCCAGACAGCTCCCAACACCTCACTAGCTCCCAACCTATCC	482
Db	636	GACCCCTACTCAGGCTCGGTACCCAGACAGCTCCCAACACCTCACTAGCTCCCAACCTATCC	695
QY	483	CACGATCTCCAGGCTCTTATATGCGCGTTTGGATACAGATGATGAAAGCAACCAATG	542
Db	696	CACGATCTCCAGGCTCTTATATGCGCGTTTGGATACAGATGATGAAAGCAACCAATG	755
QY	543	TGTGGATGTGGACGAGTGTGCAACGATATCCACAGAGTGCAACCCACCCAGATTCGAT	602
Db	756	TGTGGATGTGGACGAGTGTGCAACGATATCCACAGAGTGCAACCCACCCAGATTCGAT	815
QY	603	CAATACCTGAAGGCGGGTACACCTGCTCTCTGCAACGACAGGATATTGGCTTCTGGAAGGCCA	662
Db	816	CAATACCTGAAGGCGGGTACACCTGCTCTCTGCAACGACAGGATATTGGCTTCTGGAAGGCCA	875
QY	663	GTGGTTGACATGATGAAATGCGTATAGTGTACTCTCCAGACAGCTCTGTGGAAATGTTCC	722
Db	876	GTGGTTGACATGATGAAATGCGTATAGTGTACTCTCCAGACAGCTCTGTGGAAATGTTCC	935
QY	723	TGATTCCTATTCCTGTACATGCAAACCCGTGGTTTACCCTCAATGAGATGGAAGGCTCTG	782
Db	936	TGATTCCTATTCCTGTACATGCAAACCCGTGGTTTACCCTCAATGAGATGGAAGGCTCTG	995
QY	783	CCAGATGTGACAGAGTGTGCCACCGAGAACCCCTGCGTGCCTGCAACCTGCTCAACCTTA	842
Db	996	CCAGATGTGACAGAGTGTGCCACCGAGAACCCCTGCGTGCCTGCAACCTGCTCAACCTTA	1055
QY	843	CGGCTCTTTCATCTGCGCGTGTGACCAAGATATGAATTGAGGAAGATGGCGTTCAATG	902
Db	1056	CGGCTCTTTCATCTGCGCGTGTGACCAAGATATGAATTGAGGAAGATGGCGTTCAATG	1115
QY	903	CAGTATATGAGCAGTGTGACGCTTCTGTGATTCCTCTGCCAACATGAGTGTGAACA	962
Db	1116	CAGTATATGAGCAGTGTGACGCTTCTGTGATTCCTCTGCCAACATGAGTGTGAACA	1175
QY	963	GCCCGGACATTAATCTTGCTCTCTCCCTCCAGAGTTAATCTGCTGTGATGAGCAACCGAAG	1022
Db	1176	GCCCGGACATTAATCTTGCTCTCTCCCTCCAGAGTTAATCTGCTGTGATGAGCAACCGAAG	1235
QY	1023	CTGCGCAAGACATCAACGAATGTGTGACACAGGAACACACAGTGTCAACCTGTGACGACAGTG	1082
Db	1236	CTGCGCAAGACATCAACGAATGTGTGACACAGGAACACACAGTGTGTGACGACAGTG	1295
QY	1083	CTACAAATTTCAAGGGGGCTTCAAAATGCATCTGACCCCATCGCGTGTAGAGGCGCTTATCT	1142
Db	1296	CTACAAATTTCAAGGGGGCTTCAAAATGCATCTGACCCCATCGCGTGTAGAGGCGCTTATCT	1355
QY	1143	GAGAGTACAGTATAACCGCTGTATGTCTCGCTGAGAACCCGGCTGCGACAGACAGCC	1202
Db	1356	GAGAGTACAGTATAACCGCTGTATGTCTCGCTGAGAACCCGGCTGCGACAGACAGCC	1415
QY	1203	CTTTACCATTCCTTACCGGAGACATGACAGTGTGTGTGACAGCGCTCCCTTCCCGGTGACAT	1262

Matches 2325: Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
QY	3	CCCGGCGCTCCCGGCTCTCTCCACGACTCGTCGGCCCTCTGTGAATAAACACCC	62	
Db	216	CCCGGGGCTCTCCCGGCTCTCTCCACGACTCGTCGGCCCTCTGTGAATAAACACCC	275	
QY	63	GGGAGCCCCGAGGGCCAGAGAGAGGCGAGCTGCCGAGCTCCCGGGGGTCCGCCCG	122	
Db	276	GGGAGCCCCGAGGGCCAGAGAGAGGCGAGCTGCCGAGCTCCCGGGGGTCCGCCCG	335	
QY	123	CGAGCTTCTCTCCGCTTGGCATCTCTCTCGCGGCTTGGACATGCCAGATAAA	182	
Db	336	CGAGCTTCTCTCCGCTTGGCATCTCTCTCGCGGCTTGGACATGCCAGATAAA	395	
QY	183	AAGGATACACAGTGTACATTCGTGGCTCTGTCTTCCAGACCCCTGGGAATGCACAGC	242	
Db	396	AAGGATACACAGTGTACATTCGTGGCTCTGTCTTCCAGACCCCTGGGAATGCACAGC	455	
QY	243	ACAGTCACGAAATGGCTTTGACCTGGATCGCCAGTCAGACAGTGTATTAATATGATGA	302	
Db	456	ACAGTCACGAAATGGCTTTGACCTGGATCGCCAGTCAGACAGTGTATTAATATGATGA	515	
QY	303	ATGCCAAGACATCCCGAGGCGTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG	362	
Db	516	ATGCCAAGACATCCCGAGGCGTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG	575	
QY	363	GTATTTATGCAATCCCGGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC	422	
Db	576	GTATTTATGCAATCCCGGAGCAAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC	635	
QY	423	GACCCCTACTCAGTCCGTAACCCAGAGCTGCCCAACACTCTCAGCTCCAAATATCC	482	
Db	636	GACCCCTACTCAGTCCGTAACCCAGAGCTGCCCAACACTCTCAGCTCCAAATATCC	695	
QY	483	CAGGATCTCAGGCGCTTATATGCCGCTTGGATGATGATGATGAAGAACCAATG	542	
Db	696	CAGGATCTCAGGCGCTTATATGCCGCTTGGATGATGATGATGAAGAACCAATG	755	
QY	543	TGTGATGTGAGAGAGTGTCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	602	
Db	756	TGTGATGTGAGAGAGTGTCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	815	
QY	603	CAATACTGAAGCGGGTACACTGTCTCTGCAACCGAGGATATTTGGCTTGGGAAGCCA	662	
Db	816	CAATACTGAAGCGGGTACACTGTCTCTGCAACCGAGGATATTTGGCTTGGGAAGCCA	875	
QY	663	GTGCTTAGACATGATGAATGTGCTATGTGTACTGTCAGAGAGCTCTGTGGGAATGTCC	722	
Db	876	GTGCTTAGACATGATGAATGTGCTATGTGTACTGTCAGAGAGCTCTGTGGGAATGTCC	935	
QY	723	TGATCTCTATCTTGTACATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTG	782	
Db	936	TGATCTCTATCTTGTACATGCAACCTGTTTACCTCAATGAGATGGAAGTCTTG	995	
QY	783	CCAAGATGTGAACGAGTGTGCCACCGAAGAACCCCTGGCTGCAAACTGCTGACACCTTA	842	
Db	996	CCAAGATGTGAACGAGTGTGCCACCGAAGAACCCCTGGCTGCAAACTGCTGACACCTTA	1055	
QY	843	CGGCTTTTCATGTGCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCGTTCATTTG	902	
Db	1056	CGGCTTTTCATGTGCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCGTTCATTTG	1115	
QY	903	CAGTGATGTGAACGAGTGTCTGTGAGTTCCTGTGCAACATGATGATGTGTGAACCA	962	
Db	1116	CAGTGATGTGAACGAGTGTCTGTGAGTTCCTGTGCAACATGATGATGTGTGAACCA	1175	
QY	963	GGCCGGCACATCTTCTGTCTGTGCGCTCAAGGCTACATCTGTGTGATGAACCGAAG	1022	
Db	1176	GGCCGGCACATCTTCTGTCTGTGCGCTCAAGGCTACATCTGTGTGATGAACCGAAG	1235	
QY	1023	CTGCCAAGACATGAAGAAATGTGAGCAGACGAAGAACACAGTGCAGACGACAGCTG	1082	
Db	1236	CTGCCAAGACATGAAGAAATGTGAGCAGACGAAGAACACAGTGCAGACGACAGCTG	1295	

QY	1083	CTACAAATTTACAAAGGGGCTTCAATGATGACGACCCCATCCGCTGTGAGAGACCTTATCT	1142	
Db	1296	CTACAAATTTACAAAGGGGCTTCAATGATGACGACCCCATCCGCTGTGAGAGACCTTATCT	1355	
QY	1143	GAGATACATGATTAACCGCTGTATGTCTGTGTAGAACCTGGCTGCAGAGACACGCC	1202	
Db	1356	GAGATACATGATTAACCGCTGTATGTCTGTGTAGAACCTGGCTGCAGAGACACGCC	1415	
QY	1203	CTTTACCATCTGTACCGGGGACATGAGACGTGTGACAGACGCTCCGCTGACAT	1262	
Db	1416	CTTTACCATCTGTGTACCGGGGACATGAGACGTGTGACAGACGCTCCGCTGACAT	1475	
QY	1263	CTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCTATTACATTTTCCAGATCAAAATC	1322	
Db	1476	CTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCTATTACATTTTCCAGATCAAAATC	1535	
QY	1323	TGGGAATGAGGCGCAGAGAAATTTACATGCGGCCAAACGGGCCCATACAGTCCACCTGGT	1382	
Db	1536	TGGGAATGAGGCGCAGAGAAATTTACATGCGGCCAAACGGGCCCATACAGTCCACCTGGT	1595	
QY	1383	GATGACACGGCCCATCAAGGGGCCCGGGAATCCAGCTGAGCTGGAATGATCACTGT	1442	
Db	1596	GATGACACGGCCCATCAAGGGGCCCGGGAATCCAGCTGAGCTGGAATGATCACTGT	1655	
QY	1443	CAACACTGTATCAACTTCAGAGGACAGCTCCGTGATCCGACTGCGGATATATGTGCGCA	1502	
Db	1656	CAACACTGTATCAACTTCAGAGGACAGCTCCGTGATCCGACTGCGGATATATGTGCGCA	1715	
QY	1503	GTACCCATTTGAGCTTCGGGCTGAGGCTCCGAGCTGCGCTCTCATTTGGCACCAAGGGA	1562	
Db	1716	GTACCCATTTGAGCTTCGGGCTGAGGCTCCGAGCTGCGCTCTCATTTGGCACCAAGGGA	1775	
QY	1563	CAGGAGAAGAGGAATTAACAGAGAAATAGAGGCGACAGAGGATTAAGCATTTCCG	1622	
Db	1776	CAGGAGAAGAGGAATTAACAGAGAAATAGAGGCGACAGAGGATTAAGCATTTCCG	1835	
QY	1623	CTGAACGTTTTCCCGGAGAGTCAACCCCGACTTCTGTGACTGCACCTGACTATTTGACAGA	1682	
Db	1836	CTGAACGTTTTCCCGGAGAGTCAACCCCGACTTCTGTGACTGCACCTGACTATTTGACAGA	1895	
QY	1683	CTGTGTACCTGTGACGACTTCCGACCCCGAGTCTCTATGATACATGATTAACAAAGTAT	1742	
Db	1896	CTGTGTACCTGTGACGACTTCCGACCCCGAGTCTCTATGATACATGATTAACAAAGTAT	1955	
QY	1743	ATCATTCCTCCCGATGAGAAATGTTGGGAATTTCAAGGCGCTCATGTTATTTCCA	1802	
Db	1956	ATCATTCCTCCCGATGAGAAATGTTGGGAATTTCAAGGCGCTCATGTTATTTCCA	2015	
QY	1803	CTATTTTCAAAAGAAATAGATTAGTTTGCAGGGGCTGAGTCTATGTTCAAAAGACTGTG	1862	
Db	2016	CTATTTTCAAAAGAAATAGATTAGTTTGCAGGGGCTGAGTCTATGTTCAAAAGACTGTG	2075	
QY	1863	AACAGCTGTGTGACATCTTCCACCTCTCTCCACTCTCTCTCTCTCTCTCTCTCTCTCT	1922	
Db	2076	AACAGCTGTGTGACATCTTCTCCACTCTCTCTCTCCACTCTCTCTCTCTCTCTCTCTCTCT	2135	
QY	1923	CAAAAGACCGGGAGCTGGCGGGGAACCTGGGAGATAGTGTCTTTTCTGTAACACA	1982	
Db	2136	CAAAAGACCGGGAGCTGGCGGGGAACCTGGGAGATAGTGTCTTTTCTGTAACACA	2195	
QY	1983	GAGAAGGCTATGTAACAAACACAGCAGATCGAAGGGTTTTAGAGAAATGTGTTCAA	2042	
Db	2196	GAGAAGGCTATGTAACAAACACAGCAGATCGAAGGGTTTTAGAGAAATGTGTTCAA	2255	
QY	2043	AACCATGCTGTGATTTTCAACATTAAGAAAGTTTCAAGTTGCTTAAATTTGATTAAC	2102	
Db	2256	AACCATGCTGTGATTTTCAACATTAAGAAAGTTTCAAGTTGCTTAAATTTGATTAAC	2315	
QY	2103	GGTTTATTTCTGTCTGTCTATTTTGAATTTTAAATAAATATGTCGATGATTTCTTC	2162	
Db	2316	GGTTTATTTCTGTCTGTCTATTTTGAATTTTAAATAAATATGTCGATGATTTCTTC	2375	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 107.424 Seconds

(without alignments)
6646.008 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328

Sequence: 1 gaccgcgcgcctcccccgtg.....tgatcaatattgaagaga 2328

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2.6/ptodata/1/lna/PCPTS.COMB.seq:*
6: /cgn2.6/ptodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2324.4	99.8	2550	2	US-08-884-072-2
2	2324.4	99.8	2550	4	US-09-212-168-2
3	896	38.5	1260	3	US-09-188-930-67
4	386	16.6	1531	2	US-08-833-963C-1
5	386	16.6	2018	3	US-08-980-514-2
6	331.4	14.2	1018	3	US-09-188-930-259
7	241.4	10.4	2512	4	US-09-248-757-1
8	106.8	4.6	3373	2	US-08-897-443-2
9	89.4	3.8	401	4	US-09-643-597-278
10	80.2	3.4	220	4	US-09-404-879A-340
11	58.2	2.5	2461	1	US-08-282-141-1
12	53.2	2.3	3759	3	US-08-479-722B-3
13	52.4	2.3	5499	3	US-08-479-722B-1
14	52.4	2.3	5502	5	PCR-US95-02251-17
15	51.6	2.2	3753	5	PCR-US95-02251-2
16	51.6	2.2	4314	1	US-08-199-780-2
17	51.6	2.2	4314	1	US-08-316-650-2
18	51.6	2.2	5089	6	5177197-31
19	50.4	2.2	6412	4	US-08-652-877-17
20	50.4	2.2	6412	4	US-08-476-515A-17
21	50.4	2.2	14044	4	US-08-652-877-85
22	50.4	2.2	14044	4	US-08-652-877-89
23	50.4	2.2	14080	4	US-08-652-877-87
24	50.4	2.2	14083	4	US-08-476-515A-83
25	50.4	2.2	14086	4	US-08-652-877-83
26	45.8	2.0	7218	1	US-08-232-463-14
27	45.6	2.0	3460	2	US-08-751-305-1

28	44.8	1.9	2181	6	5208144-36	Patent No. 5208144
29	43.4	1.9	3546	4	US-08-872-757-3	Sequence 3, Appli
30	41.4	1.8	2492	4	US-09-381-779-8	Sequence 1, Appli
31	39.8	1.7	3156	4	US-09-284-819-8	Sequence 8, Appli
32	39.6	1.7	1639	2	US-08-468-819-77	Sequence 77, Appli
33	38	1.6	6048	4	US-09-634-920-3	Sequence 3, Appli
34	37.8	1.6	1611	4	US-09-249-697A-2	Sequence 2, Appli
35	37.8	1.6	1611	4	US-09-363-316B-2	Sequence 2, Appli
36	37.8	1.6	2365	4	US-09-249-697A-5	Sequence 5, Appli
37	37.8	1.6	2365	4	US-09-249-697A-18	Sequence 18, Appli
38	37.8	1.6	2365	4	US-09-363-316B-5	Sequence 23, Appli
39	37.8	1.6	2365	4	US-09-363-316B-23	Sequence 23, Appli
40	37.2	1.6	2397	6	5258288-2	Patent No. 5258288
41	37	1.6	6344	4	US-08-843-417-1	Sequence 1, Appli
42	36.6	1.6	5099	4	US-09-610-040-5	Sequence 5, Appli
43	36.6	1.6	72828	3	US-09-009-913-1	Sequence 1, Appli
44	35.4	1.5	825	1	US-08-312-870-6	Sequence 6, Appli
45	35.4	1.5	1338	1	US-08-307-444A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-884-072-2
Sequence 2, Application US/08884072
Patent No. 5872234
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884, 072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOT01
CLONE: 45517
US-08-884-072-2
Query Match 99.8%; Score 2324.4; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;

OY	664	GGCTTAGACATTTGANTGANTGTCGATGCGGTAACTGACCGCAGACGCTCTGTGCGAAATGTTCT	723
Dp	766	TGCTTAGATATTGAGAAATGTGCTATGTGTTACTGCGACGACGCTGTGGCAAAATGTTTCCA	825
OY	724	GGATTCCTATTTCTTGTACATTCGCAACCCCTGGTTTATCCCTCAATGAGAGATGAAAGCTCTTGC	783
Dp	826	GGATTCCTATTTCTTGTACATTCGCAACCCCTGGTTTATCCCTCAATGAGAGATGAAAGCTCTTGC	885
OY	784	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGAACCCCTGCTCAACACCTTAC	843
Dp	886	CAAGATGTGAACGAGTGTGCCAACTGAGAAATCCCTGTGTTCAGACCTGTGTCACAACTAT	945
OY	844	GGCTTTTCATCTGGCCGCTGTGACCCAGATATGAATTGTGGAGAAAGATGGCGTTTCATTCG	903
Dp	946	GGCTTTTCATCTGGCCGCTGTGACCCAGATATGAATTGTGGAGAAAGATGGCGTTTCATTCG	1005
OY	904	AGTATATGAGACGAGTGCAGCTTCTGATGTCCCTCCAAACATGAGTGTGTGAACCG	963
b	1006	AGTATATGAGACGAGTGCAGCTTCTGATGTCCCTCCAAACATGAGTGTGTGAACCG	1065
OY	964	CCCGGCACATACTTCTGCTCTGCCCTTCAGAGCTACATCTGCTGATGAGACACGGAAGC	1023
Dp	1066	CCCGGCCTATACTTCTGCTCTGCCCTTCAGAGCTACATCTGCTGATGAGACACGGAAGC	1125
OY	1024	TGCCAAGACATCAACGAATGTGAGCACAGGAACACAGCTGCACCTGACAGACGTGC	1083
Dp	1126	TGCCAAGATATCAATGAATGTGAGCACCGGAACACAGCTGTACTTCACTGACGCTTGC	1185
OY	1084	TACATATTACAAGGGGGGCTTCAAATGCAATGCATGCCCATCCGCTGTGAGAGACCTTATCTG	1143
Dp	1186	TACATATTACAAGGGGGGCTTCAAATGCAATGCATGCCCATCCGCTGTGAGAGACCTTATCTG	1245
OY	1144	AGGATCACTGATTAACCCGCTGTATGTGTCCTGCTGAGAACCCCTGCGTGCAGAGACAGCC	1203
Dp	1246	CTGATTTGGTGAACAAACCGCTGTATGTGTCCTGCTGAGCACACAGCTGTGAGAGACCTTATCTG	1305
OY	1204	TTTACCATCTTGTACCGGGACATGAGAGCTGGTGTGAGAGCCTCCGTTCCCGCTGCATCTC	1263
Dp	1306	TTTACCATCTTGTATCCGGACATGGATGTGTGTCTGAGGACCCCTCCGTTCCCGCTGCATCTC	1365
OY	1264	TTTCAAAATGCAAGGCACGACCCGGATCCCTGGGGGCTTATTCATTTTCCAGATCAAAATCT	1323
Dp	1366	TTTCCAGATGCAAGGCACCAACCCCGATACCCCTGTGCTCTTTCATTTTCCAGATCAAAATCT	1425
OY	1324	GGGATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCCACACCTGGTG	1383
Dp	1426	GGCAACGAGGGTCTGAGAGTTCTTATATGCGCAAAAGGGCCCTATACAGTCCACACCTGGTG	1485
OY	1384	ATGACAGCCCCCATATCAAAGGGCCCCGGGAAATCCAGCTGGAATTGGAATGATCATCTGTC	1443
Dp	1486	ATGACAGCCCCCATATCAAAGGGCCCCGGGAAATCCAGCTGGAATTGGAATGATCATCTGTC	1545
OY	1444	AACACTGCATCAACTTCAGAGGGACGTCGCGTGAATCCGATGCGAGATATATGTCTCCGAG	1503
Dp	1546	AACACTGCATCAACTTCAGAGGGACGTCGCGTGAATCCGATGCGAGATATATGTCTCCGAG	1605
OY	1504	TACCATTTCTGAGCTCGGCGTGGAGCCTCGAGAGCTCCCTGTATTGGACCAAGGGAC	1563
Dp	1606	TATCCGTTCTAGGCTCTGCGGTAAGGCTCTGACATCTGCTTTTACACGACCCGAGGGAC	1665
OY	1564	AGGAGAAAGAGAAATTAACAGAGAGATGAGACGACACAGACGTTAAGCATTTTCTGTC	1623
Dp	1666	GGGAGGAGAAAGGAAAC--CAGCAAGAAATGAGACGAGACAGACATTGCACTTTCCTGTC	1723
OY	1624	TGAACGTTTCCCCGAAAGATGACGCCCGGACTTCTGACTCTACCTGTACTATTGGAGAC	1683
Dp	1724	TGAATATCTCTCTGGGGGATATGAGCTAG-CATCTTGGACCCATTATCTGTACTATTGGAGA-	1781
OY	1684	CTGACACCCCTGAGAGCTTGGCACCCCGGATGCTTATGATATCACTTATCAAAAAGTATTA	1743
Dp	1782	TGCTTACTCTTAAGAGACACCTCGCTCTGATGTTCTTATGATGACAGTTATCCAAAAGTGTTC	1841

QY	1744	TCATGGCTCCCGCTGATACAAATATGGTGGTAATTTTCAGAGCCCTCAGTTATTTCCAC	1803
Db	1842	ATCTT-AGCCCCGTGATATGAGGTTGCCAGTACCTCTTCAAAGCCCTTCATTATTTTCAT	1900
QY	1804	TATTTT----CAAGAAATAATAGATTAGGTTTGCGGGGGGTCTGAG---TCATATGTTCAAG	1856
Db	1901	CGTTTTTATAAAAAAGAAATAATAGATTAGATTGCTGGGGTATAGAGTCCTCGAANGTCAAA	1960
QY	1857	ACTGTGAACAGCTTGTCTGCTCACTTCTTCACCTCTTCCACTCTCTCTCACTGTGTTACT	1916
Db	1961	AGACGAGATGGCTTGCCTCCACACCTTCCTCTCTCTC-CTCCATCTCTTGGTCGATTGCT	2019
QY	1917	GCTTTGCAAAAGACCCGG--GAGCTGGCGGGGAGACCCGTGGATAGCTATGTTGCTTTTG	1974
Db	2020	GCTTTGCAAAAGTCTCTCATGGGCTCTGTGGAAATGCTGGAAATAGCTATGTTGCTTTTG	2079
QY	1975	CGTACACAGAGAAGGCTATGTAAACAAACACACAGAGAGTGAAGGGTTTTTTAGAGATG	2034
Db	2080	CATGTTCTGAGAAAGGCTATGGGAACACACACAGAGAGATCGAA-GGTTTTTATNAGATC	2138
QY	2035	TGTTTCAAAACCATGCTCTGTATTTTCAACCATTAAGAGATTTCAGTTGCTTTAAATT	2094
Db	2139	TATTTTAAAAACATCTGTGTTATTTTTCAGCTAAAGAAATTTTATGTTGCTTTAAATT	2198
QY	2095	TGTTAAACGGTTATTTCTGTCTGTTCATTTTATAGTATTTTAAAAAATATGTGCGTAA	2154
Db	2199	TGTTATGATGTGTTTAAACCTTTTCTTATTTTATTTTATGAGGCTTTTAA-----GTGGTAGA	2252
QY	2155	ATTCCTTCGAAAGGCTTCACAGACATGCTATGTCTCTCTTCCAAACCCAGCTCTCTC	2214
Db	2253	ATTCCTTCGAAAGGCTC-TCAATATCAATTTATGTTTCAGTCTTTCCAAACCTCATGCTTTCC	2311
QY	2215	TTCATTTTAGCCACAGTGTTCCTTTGAGAGACCCCTTAATCTTGTCTTTTGAATTTT	2274
Db	2312	TGCATCTTAGCCCGAGTTTTTAC---GAAGACCCCTTAATCATGCTTTTATTAAGAGTTTTT	2368
QY	2275	ACCCAATTGAATGGAAATGACAGAGTCTCCAAATGATTAATATTTGGAGAGA	2328
Db	2369	ACCCAAGCTGCGTTGGAAACAGAGGATATCCAGACTGATTAATATATTTGAACAAA	2422

Search completed: July 3, 2003, 11:59:30
Job time : 522.747 secs

```

QY 1504 TACCATTCTAGAGCTGGGGCTGGAGCCCTCCAGCGTCTCTCATTTGGACCAAGGAC 1563
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1606 TATCCGTTCTGAGAGCTTGCTGGCTGTGACACTCTTTCACAGCAGCAGGAGAC 1665
QY 1564 AGAGAGAGAGAGAGATTAACAGAGATGAGAGAGACAGACAGCTTAGGCTTCTGCG 1623
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1666 GGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1723
QY 1624 TGAAGCTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1683
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1724 TGAATTTCTCTGGGGGATCAGCTTAG-CACTTGGCCATCTGATCTATTTGAGAG- 1781
QY 1684 CTGTACACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1743
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1782 TGGTCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1841
QY 1744 TCATTGCTCCCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1803
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1842 ATCTT-AGCCCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1900
QY 1804 TATTTT-----CAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1856
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1901 CGTTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1960
QY 1857 ACTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1916
  || || || || || || || || || || || || || || || || || || || || || ||
Db 1961 AGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2019
QY 1917 GCCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1974
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2020 GCCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2079
QY 1975 CGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2080 CAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2138
QY 2035 TGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2094
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2139 TATTTTAAATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2198
QY 2095 TGTATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2154
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2199 TGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2252
QY 2155 ATTCCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2214
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2253 ATTCCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2311
QY 2215 TCCATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2274
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2312 TGCATCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2368
QY 2275 ACCCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2328
  || || || || || || || || || || || || || || || || || || || || || ||
Db 2369 ACCCACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2422

```

RESULT 15

AAZ40031

AAZ40031 standard; DNA: 2429 BP.

AAZ40031;

15-FEB-2000 (first entry)

Full length mouse A55 protein coding sequence.

A55 protein; mouse; smooth muscle proliferation; tissue generation;
 vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
 vascular endothelial thickening; haematopoietic cell-regulation; cytokine;
 percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
 actin; inhibitor; chemotaxis; thrombosis; cadherin; therapy;
 tumour metastasis inhibitor; ss.

```

XX XX Mus musculus.
OS OS
PN PN MO9955864-A1.
PD PD 04-NOV-1999.
PE PE 28-APR-1999: 99WC-JP02284.
PR PR 28-APR-1998: 98JP-0119731.
XX XX (ONOX) ONO PHARM CO LTD.
PA PA Honjo T, Tashiro K, Nakamura T;
PI PI WPI: 2000-038647/03.
DR DR P-P-SDB; AAY54991.
PT PT Novel human polypeptides for treatment of, e.g. arteriosclerosis and
  myoma -
XX XX Example 5: Page 64-68; 87pp; Japanese.
CC CC This sequence encodes the mouse A55 protein. The invention relates to the
CC CC human A55 protein. The protein can be used for the treatment of diseases
CC CC due to abnormal proliferation of smooth muscle. The polypeptides can be
CC CC used according to their inhibition of the proliferation of vascular smooth
CC CC muscle cells, particularly in treating arteriosclerosis or re-narrowing
CC CC by vascular endothelial thickening after percutaneous transluminal
CC CC coronary angioplasty (PTCA), or myoma, haematopoietic cell-regulatory
CC CC activity, cytokine activity, tissue generation/repair activity, blood
CC CC coagulation/thrombotic activity, receptor/ligand activity,
CC CC cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
CC CC a nutrient.
SQ SQ Sequence 2429 BP; 610 A; 611 C; 562 G; 645 T; 1 other;
XX XX
Query Match 62.3%; Score 1449.8; DB 21; Length 2429;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1795; Conservative 0; Mismatches 333; Indels 26; Gaps 12;
QY 184 AGGATCTACCTGTTACCATTTGCTCTGCTCTTCCAGCCCTGGAGATCCAGAGCA 243
  || || || || || || || || || || || || || || || || || || || || || ||
Db 286 ATGATCTACCTGTTACCATTTGCTCTGCTCTTCCAGCCCTGGAGATCCAGAGCA 345
QY 244 CAGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
  || || || || || || || || || || || || || || || || || || || || || ||
Db 346 CAGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405
  || || || || || || || || || || || || || || || || || || || || || ||
QY 304 TGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
  || || || || || || || || || || || || || || || || || || || || || ||
Db 406 TGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
  || || || || || || || || || || || || || || || || || || || || || ||
QY 364 TATTTATGATTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
  || || || || || || || || || || || || || || || || || || || || || ||
Db 466 TATTTATGATTTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
QY 424 ACCCCCTACCTGAGTCCGTTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 483
  || || || || || || || || || || || || || || || || || || || || || ||
Db 526 ACATCTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
  || || || || || || || || || || || || || || || || || || || || || ||
QY 484 ACATCTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
  || || || || || || || || || || || || || || || || || || || || || ||
Db 586 ACATCTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 645
  || || || || || || || || || || || || || || || || || || || || || ||
QY 544 GTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
  || || || || || || || || || || || || || || || || || || || || || ||
Db 646 GTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 705
  || || || || || || || || || || || || || || || || || || || || || ||
QY 604 AATACCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
  || || || || || || || || || || || || || || || || || || || || || ||
Db 706 AACACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765

```



```

Db      1099  GAGACAGGCATTCACCATCTCTGATCGGAGACATGATGTCGTCAGAGACGTCCTCCGCTC 1158
OY      1253  CCGGTGACATCTTCCAAATGCAAGCCGACCCGCTACCTGGGGCCCTATTACATTTTCC 1312
Db      1159  CTGCTGACATCTTCCAGATGCAACCAACCCGATCCCTGGGCTTATTACATTTTCC 1218
OY      1313  AGATCAATCTGGGAATGAGGAGAAATTTTACATCGGCAACAGGGCCCATCAGTG 1372
Db      1219  AGATCAATCTGGCAACGAGGGTCGAGAGTTCATATGCGGCAACAGGGCCCTATCAGTG 1278
OY      1373  CCACCTGGTATGACAGCCCATCAAGGGCCCGGGAATTCAGCTGGACTTGGAAA 1432
Db      1279  CCACCTGGTATGACAGCCCATCAAGGGCCCGGGAATTCAGCTGGACTTGGAGA 1338
OY      1433  TGTATCTGTCAACATCTGTATCACTTCAGAGGACCTCGGTATCCGATCGGGATAT 1492
Db      1339  TGTATCTGTCAACATCTGTATCACTTCAGAGGACCTCGGTATCCGATCGGGATAT 1398
Y      1493  ATGTGTCGAGTACCATCTGTGAGCCCTGGGCTGAGACCTCCGACGCTCCTCATTTGG 1552
Db      1399  ATGTGTCGAGTATCCGCTTGTAGCCCTGTGAGGCTCTGTACACTGCTTTCACGAG 1458
OY      1553  CACCAAGGACAGAGAGAGAGAAATTAACAGAGATGAGACGACACAGACGTTAG 1612
Db      1459  CACCGAGGAGCGGAGAGAGAGAA--ACCGACAGAAATGAGAGCGAGACGATTCG 1516
OY      1613  GATTTCCGTCGATGAGTTTCCCGAAGATGACAGCCCGACTCTGCTCTCATCTGTA 1672
Db      1517  ACCCTTCCGTCGATGATCTCTGGGGCATCAGCCCTAG-CATTTGACCATATCTGTA 1575
OY      1673  CTATTCGACAGCTGTACCCCTGAGAGACTTCCACCCGATTCCTATGATACAGTATC 1732
Db      1576  CTATTCGACA-TGCTCACTCTGAAGACACCCCTGCTCAGTTCTATGATGAGATATC 1634
OY      1733  AAAAGTATATCATCTGCTCCCTGATAGAGATTTGGTAAATTTTCAAGCCCTTGA 1792
Db      1635  CAAAGTGTTCATCTT-AGCCCTGATATGAGTTGGCAGACTCTTCAAGCCCTTCA 1693
OY      1793  TTTATTTCCACTTTT---CAAGAAATATGATTTGTTGGGGGCTGTGACTT-- 1846
Db      1694  TTTATTTCCACTTTTATATAAAGAAATATGATTTGTTGGGGTATGATCTCTC 1753
OY      1847  -ATGTCAAGAGACTGTGAACAGCTGTGCTACATCTTACCTTCACCTCTCTCTCTC 1905
Db      1754  GAAAGTCAAAAGACCTGAGTGGCTGTGCTCTCACCTCTCTCTCTCTCTCTCTCT 1812
OY      1906  ACTGTGTACTGCTTTGGCAA--GACCCGGAGCTGGCGGGAAACCTGGAGATGCTAG 1963
Db      1813  GCTGCTGCTGCTTTGGCAAAGTCTCATGGGCTCTGGGAAATGCTGGAAATGCTAG 1872
OY      1964  TTTGCTTTTGGCTACAGAGAGAGGCTATGTAACAAACACAGAGATGCAAGGGTT 2023
Db      1873  TTTGCTTTTGGCTACAGAGAGGCTATGTAACAAACACAGAGATGCAAGGGTT 1931
OY      2024  TTTAGAGAAATGCTTTCAAAACCATGCTGTATTTTCAACCATAAAGAAATTTCACT 2083
Db      1932  TTTATGAGTCTTTTAAATCAACATCTGTATTTTCAAGATAAAGAAATTTAGTGTG 1991
OY      2084  GTCTTAAATTTGATTAACGCTTATTTGCTTGTCTTGTATTTGATTTTAAATAA 2143
Db      1992  TCTTTAAATTTGATTAAGGCTTAACTTTTCTTATTTTCAATTTGAGGCTTCTTAA-- 2048
OY      2144  TATGTCGTAGAAATCTCTGAAAGGCTTACAGACATGCTATGCTCTTCCCAAC 2203
Db      2049  --GAGAGAAATTTCTTCCAAAGGCC-TCAGATATATGTTATGTTCTTCCCAAC 2104
OY      2204  CGAGTCTCTCTCATTTTGGCCAGAGTCTTCTTGGAGACCCCTTATCTGCTTCT 2263
Db      2105  TCACTCTTCTCTCATCTTACCCAGTCTTTTAC--GAGAGCCCTTATATATGCTTCT 2161
OY      2264  TTAGAATTTTACCATTTGATTTGATGATGAGAGTCTCCAACTGATTTAAATTTTGA 2323

```

```

Db      2162  TAAGACTTTTACCACACTGCTTGAAGACAGAGATATCCAGACTGATTAATTTGA 2221
OY      2324  AGAGA 2328
Db      2222  AGAAA 2226

RESULT 13
AAZ40029
ID  AAZ40029 standard; DNA; 2233 BP.
XX
XX  AAZ40029;
AC  15-FEB-2000 (first entry)
DT  15-FEB-2000 (first entry)
XX
DE  Full length mouse A55 protein coding sequence.
XX
XX  A55 protein; mouse; smooth muscle proliferation; tissue generation;
KW  vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
KW  vascular endothelial thickening; haematopoietic cell-regulator; cytokine;
KW  percutaneous transmural coronary angioplasty; blood coagulation; PTCA;
KW  actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
KW  tumour metastasis inhibitor; ss.
XX
XX  Mus musculus.
OS  Mus musculus.
XX  MO9955864-AI.
PN  MO9955864-AI.
XX
XX  04-NOV-1999.
PD  04-NOV-1999.
XX
XX  28-APR-1999; 99MO-JP02284.
PF  28-APR-1999; 99MO-JP02284.
XX
XX  28-APR-1998; 98JP-0119731.
PR  28-APR-1998; 98JP-0119731.
XX
XX  (ONOR ) ONO PHARM CO LTD.
PA  (ONOR ) ONO PHARM CO LTD.
XX
XX  Honjo T, Tashiro K, Nakamura T;
PI  Honjo T, Tashiro K, Nakamura T;
XX
XX  WPI: 2000-038647/03.
DR  WPI: 2000-038647/03.
XX
XX  P-PSDB; AAT54990.
DR  P-PSDB; AAT54990.
XX
XX  Novel human polypeptides for treatment of, e.g. arteriosclerosis and
PT  myoma -
XX
XX  Example 4; Page 52-55; 87pp; Japanese.
XX
XX  This sequence encodes the mouse A55 protein. The invention relates to the
CC  human A55 protein. The protein can be used for the treatment of diseases
CC  due to abnormal proliferation of smooth muscle. The polypeptides can be
CC  used according to their inhibition of the proliferation of vascular smooth
CC  muscle cells, particularly in treating arteriosclerosis or re-narrowing
CC  by vascular endothelial thickening after percutaneous transmural
CC  coronary angioplasty (PTCA), or myoma, haematopoietic cell-regulatory
CC  activity, cytokine activity, tissue generation/repair activity,
CC  actin/inhibin activity, taxis and chemotaxis activity, blood
CC  coagulation/thrombotic activity, receptor/ligand activity,
CC  cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
CC  a nutrient.
XX
XX  Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other:
SQ

```

Query Match 63.6%; Score 1480.8; DB 21; Length 2233;
 Best Local Similarity 82.7%; Pred. No. 0;
 Matches 1841; Conservative 0; Mismatches 358; Indels 26; Gaps 12;

```

OY      113  GTCCCGCCGAGCTTTCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 172
Db      19  GTCCACCCGAGAGCGCTTCCTCCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 78
OY      173  CAGGATTAATAAGATATCTACTGTTTACATTTGCTGCTCTCTCTTCCAGCCCTGGGA 232
Db      79  CAGGATTAATAAGATATCTACTGTTTACATTTGCTGCTTCGCTTCGCTTCGCTTCG 138

```


QY 2170 CTTCAGACATGCTATGTTCTGTTCCCAACCCAGTCCTCCATTTAGCCAG 2229
 |||||
 Db 1381 CTTCAGACATGCTATGTTCTGTTCCCAACCCAGTCCTCCATTTAGCCAG 1440
 QY 2230 TGTTCCTTTGAGAGCCCTTAATCTGCTTTTGTAGAAATTTTACCCTGATG 2289
 |||||
 Db 1441 TGTTCCTTTGAGAGCCCTTAATCTGCTTTTGTAGAAATTTTACCCTGATG 1500
 QY 2290 AATGAGAGGCTCCCAACGATTAATTAATTTGAAGAG 2327
 |||||
 Db 1501 AATGAGAGGCTCCCAACGATTAATTAATTTGAAGAG 1538

RESULT 12

AAZ39384

AAZ39384 standard; DNA; 2233 BP.

AAZ39384;

22-FEB-2000 (first entry)

Smooth muscle proliferation modulating protein encoding DNA.

Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 endothelial thickening; percutaneous transluminal coronary angioplasty;
 myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 metastasis; nutrient; ss.

Mus musculus.

MO9955863-A1.

04-NOV-1999.

28-APR-1999; 99MO-JP02283.

28-APR-1998; 98JP-0119731.

(ONOR) ONO PHARM CO LTD.

Honjo T, Tashiro K, Nakamura T;

WPI: 2000-038646/03.

P-PDB: AAY56750.

Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

Claim 5; Page 47-50; 70pp; Japanese.

The invention provides mouse polypeptides for treatment of diseases due
 to abnormal proliferation of smooth muscle. The polypeptides can be
 produced by standard recombinant methodology. The polypeptides can be
 used according to their inhibition of the proliferation of vascular
 smooth muscle cells, particularly in treating arteriosclerosis or re-
 narrowing by vascular endothelial thickening after percutaneous
 transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 regulatory activity, cytokine activity, tissue generation/repairation
 activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
 coagulation/thrombotic activity, receptor/ligand activity, cadherin/
 tumor metastasis inhibiting activity; tumor inhibition, and as nutrient.
 The present sequence represents a DNA encoding the protein of the
 invention which can be used for modulating smooth muscle cell
 proliferation.

Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other;

Query Match 63.6%; Score 1480.8; DB 21; Length 2233;

Best Local Similarity 82.7%; Pred. No. 0; Mismatches 358; Indels 26; Gaps 12;

Matches 1841; Conservative 0;

QY 113 GTCCGCCGCGAGCTTCTTCTGCGCTTCGATCTCTCTCGCGCGCTTGGACATGC 172
 |||||
 Db 19 GTCCACCGGAGACGCTTCTCTCTCGCGCTTCTCTCGCGCGCTTGGATATGC 78
 QY 173 CAGGAATTAAGAGATCTACTGTTTACCATTTGCTGCTCTGTTCCAGCCCTGGGA 232
 |||||
 Db 79 CAGGATTAAGAGATCTACTGTTTACCATTTGCTGCTCTGTTCCAGCTGGGA 138
 QY 233 ATGCAGAGGACAGTGCAGATGCTTGCCTGATGCGGAGGACAGATGTTTATG 292
 |||||
 Db 139 ATGCAGAGGACAGTGCAGATGCTTGCCTGATGCGGAGGACAGATGTTTATG 198
 QY 293 ATATTGATGATGCGGAGATCCCGAGGCTGCGGAGGACATGATGTTTATG 352
 |||||
 Db 199 ATATTGATGATGCGGAGATCCCGAGGCTGCGGAGGACATGATGTTTATG 258
 QY 353 AAAATGGCGGATTTTATGCAATTCCTGCGGACAAACCTGTTGATGAGGCGCTACTGGA 412
 |||||
 Db 259 AGAATGGCGGATTTTATGCAATTCCTGCGGACAAACCTGTTGATGAGGCGCTACTGGA 318
 QY 413 ACCCTTCTGAGCCCTGCTGAGGCTGCTGAGGCTGCGGAGGACATGATGTTTATG 472
 |||||
 Db 319 ATCCCTTCTGATCTCTCTCTGAGGCTGCTGAGGCTGCGGAGGACATGATGTTTATG 378
 QY 473 CAACCTATCCGACGATCTCCAGGCTCTTATGCGGCTTTGGATACGATGATGATAA 532
 |||||
 Db 379 CAACCTATCCGACGATCTCCAGGCTCTTATGCGGCTTTGGATACGATGATGATAA 438
 QY 533 GCAACCAATGTTGAGATGAGACGAGTGCACAGATTCACACAGTCCAAACCCAGCC 592
 |||||
 Db 439 GCAACCAATGTTGAGATGAGACGAGTGCACAGATTCACACAGTCCAAACCCAGCC 498
 QY 593 AGATCTGATCAATACGTAAGCGGCTGACCTGCTGCTGCGGACGCGGATTTGGCTTC 652
 |||||
 Db 499 AGATCTGATCAATACGTAAGCGGCTGACCTGCTGCTGCGGACGCGGATTTGGCTTC 558
 QY 653 TGAAGGCGAGTGTATGACATTTGATGATGCTGATGCTTACTGCGAGCTCTGTG 712
 |||||
 Db 559 TGAAGGCGAGTGTATGACATTTGATGATGCTGATGCTTACTGCGAGCTCTGTG 618
 QY 713 CGAATGTTCTGATCTTATTTTGTACATGCAACCTGCTTATCCCTCAATGAGATG 772
 |||||
 Db 619 CGAATGTTCTGATCTTATTTTGTACATGCAACCTGCTTATCCCTCAATGAGATG 678
 QY 773 GAAAGTCTGCGCAAGATGTAAGAGAGTGGCCAGCAGGACCTGCGCAACCTGGG 832
 |||||
 Db 679 GAAAGTCTGCGCAAGATGTAAGAGAGTGGCCAGCAGGACCTGCGCAACCTGGG 738
 QY 833 TCAACACCTAGCGCTTCTTCAATCTGCGCTGAGCCAGATATGAACTTGAAGAGATG 892
 |||||
 Db 739 TCAACACCTAGCGCTTCTTCAATCTGCGCTGAGCCAGATATGAACTTGAAGAGATG 798
 QY 893 GCGTTCAATGAGTATGAGACAGTGCAGCTTCTGAGTTCCTGCTGCAACATGAGT 952
 |||||
 Db 799 GCAATTCATGAGTATGAGACAGTGCAGCTTCTGAGTTCCTGCTGCAACATGAGT 858
 QY 953 GTGGAACACCGCGGCGCATATCTGCTGCTGCGCTGCGGACGATATGCTGCTGATG 1012
 |||||
 Db 859 GTGGAACACCGCGGCGCATATCTGCTGCTGCGCTGCGGACGATATGCTGCTGATG 918
 QY 1013 ACAACCGAGCTGCGCAAGATCAACGATGAGACAGAGAACACAGAGTGCACCTGC 1072
 |||||
 Db 919 ATACCGAGCTGCGCAAGATCAACGATGAGAGACAGAGAACACAGAGTGCACCTGC 978
 QY 1073 AGCAGAGCTGCTAATTTAAGAGGCTTCAATGATGACAGCCATCCGCTGTGAGG 1132
 |||||
 Db 979 TGCAGAGCTGCTAATTTAAGAGGCTTCAATGATGACAGCCATCCGCTGTGAGG 1038
 QY 1133 AGCCTTATCTGAGAGTATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192
 |||||
 Db 1039 AGCCTTATCTGAGAGTATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098
 QY 1193 GAGACAGCCCTTTACATCTTGTACCGGACATGAGCTGTGTACAGAGCTCTCGTTC 1252

AAK94129
ID AAK94129 standard: cDNA, 1538 BP.
AC AAK94129;
XX
DI 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2627.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KM
XX Homo sapiens.
OS
XX EPI130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PE
XX
XX 08-JUL-1999; 98JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
DR P-PSDB; AAM93220.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 2627; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
CC
XX
SQ Sequence 1538 BP; 387 A; 399 C; 350 G; 402 T; 0 other;

Query Match 65.9%; Score 1534.8; DB 22; Length 1538;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 790 GTGAACGAGTGTGCCACCGAAGACCCCTGGTGCACAACTGCGTCAACACCTAGCGCTCT 849
DB 1 GTGAACGAGTGTGCCACCGAAGACCCCTGGTGCACAACTGCGTCAACACCTAGCGCTCT 60
QY 850 TTTCATCTGCGGCTGTACCCAGATATGAACTTGAGAAATGCGCTTCATTTGAGTGTAT 909
DB 61 TTTCATCTGCGGCTGTACCCAGATATGAACTTGAGAAATGCGCTTCATTTGAGTGTAT 120
QY 910 ATGGACGAGTGTGAGCTTCTGTAGTTCCTGTGCAATGAGTGTGTGAACACCGCCGCG 969
DB 121 ATGGACGAGTGTGAGCTTCTGTAGTTCCTGTGCAATGAGTGTGTGAACACCGCCGCG 180
QY 970 ACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGATGACCAACGAGCTGCCAA 1029
DB 181 ACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGATGACCAACGAGCTGCCAA 240
QY 1030 GACATCAACGAGTGTGAGCAGAGAACCAACGTCGACCACTGACGAGACGCTGACAT 1089

DB 241 GACATCAACGAGTGTGAGCAGAGAACCAACGTCGACCACTGACGAGACGCTGCTACAT 300
QY 1090 TTACAAAGGGGCTTCAAAATGCATCGACCCCATCGGCTGTGAGAGGCTTATCGAGATC 1149
DB 301 TTACAAAGGGGCTTCAAAATGCATCGACCCCATCGGCTGTGAGAGGCTTATCGAGATC 360
QY 1150 AGTGATTAACCGCTGTATGTCTGCTGTGAGAACCCCTGCTGCAGAGACAGCCCTTTACC 1209
DB 361 AGTGATTAACCGCTGTATGTCTGCTGTGAGAACCCCTGCTGCAGAGACAGCCCTTTACC 420
QY 1210 ATCTTGTAACCGGAGACATGACGTGTGTGACGACCTCGTTCCTCCGCTGACATCTTCCA 1269
DB 421 ATCTTGTAACCGGAGACATGACGTGTGTGACGACCTCGTTCCTCCGCTGACATCTTCCA 480
QY 1270 ATGCAAGCCAGCAACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAATCTGGGAT 1329
DB 481 ATGCAAGCCAGCAACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAATCTGGGAT 540
QY 1330 GAGGCGAGAGAAATTTTACATGCGGCAACGCGGCCCATCATGAGTCCACCTGTGATGACA 1389
DB 541 GAGGCGAGAGAAATTTTACATGCGGCAACGCGGCCCATCATGAGTCCACCTGTGATGACA 600
QY 1390 CGCCCATTCAAAAGGCGCCGGAATTCACAGCTGAGCTTGGAAATGATCAGTCAACACT 1449
DB 601 CGCCCATTCAAAAGGCGCCGGAATTCACAGCTGAGCTTGGAAATGATCAGTCAACACT 660
QY 1450 GTCATCACTTCAAGAGGACGCTCGTGTATCCAGTGGAGATATATGTGTGCACTATACCA 1509
DB 661 GTCATCACTTCAAGAGGACGCTCGTGTATCCAGTGGAGATATATGTGTGCACTATACCA 720
QY 1510 TTCTGAGCCTCGGGCTGTGAGCCTTCGACGCTGCTCTCATTTGGCACAAGGACAGAGA 1569
DB 721 TTCTGAGCCTCGGGCTGTGAGCCTTCGACGCTGCTCTCATTTGGCACAAGGACAGAGA 780
QY 1570 AGAGGGAATTAACAAGAGAAATGAGAGGACAGACAGAGTATGAGCTTTCTGCTGAACG 1629
DB 781 AGAGGGAATTAACAAGAGAAATGAGAGGACAGACAGAGTATGAGCTTTCTGCTGAACG 840
QY 1630 TTTTCCCGGAAGATGACCCCGGCTTCTGACTCTCACTGCTACTATTTGACAGACTGTCA 1689
DB 841 TTTTCCCGGAAGATGACCCCGGCTTCTGACTCTCACTGCTACTATTTGACAGACTGTCA 900
QY 1690 CCTGTGAGACTTGTGCAACCCCGGCTTCTGACTCTCACTGCTACTATTTGACAGACTGTCA 1749
DB 901 CCTGTGAGACTTGTGCAACCCCGGCTTCTGACTCTCACTGCTACTATTTGACAGACTGTCA 960
QY 1750 CTCCCTGATGAGAAATGTTGTGGAATTTTCAAGGCTTCACTTATTTGACATTTT 1809
DB 961 CTCCCTGATGAGAAATGTTGTGGAATTTTCAAGGCTTCACTTATTTGACATTTT 1020
QY 1810 CAAGAAGAAATAGATTAGTTTGGGGGCTGTGAGTGTATGTTCAAGAGCTGTGAACACT 1869
DB 1021 CAAGAAGAAATAGATTAGTTTGGGGGCTGTGAGTGTATGTTCAAGAGCTGTGAACACT 1080
QY 1870 TGTGTGACCTTCTTCACTTCTTCACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1929
DB 1081 TGTGTGACCTTCTTCACTTCTTCACTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1930 CCGGAGAGTGTGCGGGGGAACCCCTGGAGTAGTACTTGTCTTTTGGCTGACACAGAGAAAG 1989
DB 1141 CCGGAGAGTGTGCGGGGGAACCCCTGGAGTAGTACTTGTCTTTTGGCTGACACAGAGAAAG 1200
QY 1990 CTATGTAAACAAACACAGCAGATGCAAGGCTTTTGAAGAAATGTGTTTCAAAACCATG 2049
DB 1201 CTATGTAAACAAACACAGCAGATGCAAGGCTTTTGAAGAAATGTGTTTCAAAACCATG 1260
QY 2050 CCGTGTATTTTCAACCAATAAAGAGTTGCTGCTTAAATTTGTATAACGGTTTAA 2109
DB 1261 CCGTGTATTTTCAACCAATAAAGAGTTGCTGCTTAAATTTGTATAACGGTTTAA 1320
QY 2110 TTCTGCTTGTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 2169
DB 1321 TTCTGCTTGTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1380

DR P-PSDB: AAU75494.
XX New nucleic acid encoding human extracellular/epidermal growth factor,
PT useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
PT also related polypeptides -
XX
XX
PS Claim 4; Fig 1; 22pp; English.
XX The invention relates to a novel polynucleotide which is at
CC least 95% identical with a sequence (ATCC 97285) encoding mature human
CC extracellular protein-1like/epidermal growth factor (EGF)-like protein,
CC EGF. Also included are the EGF EGF domains, a vector containing
CC the polynucleotide, a host cell containing the vector, anti-EGF
CC antibodies and antagonists of EGF. The polynucleotide is used for
CC recombinant production of EGF, in gene therapy, as hybridisation probes,
CC as antisense antagonists and for chromosome identification. The protein
CC is used to treat patients who require EGF, to identify specific
CC antagonists, used to treat conditions that require inhibition of EGF
CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
CC healing, neurological trauma, acquired immunodeficiency syndrome
CC (AIDS)-related dementia, ocular disorders, kidney disorders, liver
CC disorders, hair follicle growth promotion, burns, ulcers, corneal
CC incisions, corneal inflammation, neoplasms and psoriasis), to raise
CC specific antibodies and to characterise receptors. The present
CC sequence encodes EGF.
XX
SQ Sequence 1720 BP: 396 A; 515 C; 434 G; 372 T; 3 other:
Query Match 71.9%; Score 1674.2; DB 24; Length 1720;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1673; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 CCCGGCCCTCTCCCGCTGCTCTCCAGACGCTGCGCCCTCTGGAATAAACCC 62
DB 45 CCGGGCCCTCTCCCGCTGCTCTCCAGACGCTGCGCCCTCTGGAATAAACCC 104
QY 63 GCGAGCCCCGAGGCCCCAGAGAGGCGCAGCTGCCGAGCTCTCCGGGGTCCCGCC 122
DB 105 GCGAGCCCCGAGGCCCCAGAGAGGCGCAGCTGCCGAGCTCTCCGGGGTCCCGCC 164
QY 123 CGAGCTTCTCTGCGCTTCCGATCTCTCCGCGGCGCTTGGACATCCAGGAATAA 182
DB 165 CGAGCTTCTCTGCGCTTCCGATCTCTCCGCGGCGCTTGGACATCCAGGAATAA 224
QY 183 AAGGATCTCACTGTATACATTTGCTCTGCTTCCAAAGCCCTGGGAATCACAGC 242
DB 225 AAGGATCTCACTGTATACATTTGCTCTGCTTCCAAAGCCCTGGGAATCACAGC 284
QY 243 ACAGTCACGAGATGGCTTGAACCTGATGCGCAGTCAGAGACAGTGTATGATGTA 302
DB 285 ACAGTCACGAGATGGCTTGAACCTGATGCGCAGTCAGAGACAGTGTATGATGTA 344
QY 303 ATGCGGAACCATCCCGAGGCGCTGCCGAGGAGACATGATGTGTAACCAAAATGGCG 362
DB 345 ATGCGGAACCATCCCGAGGCGCTGCCGAGGAGACATGATGTGTAACCAAAATGGCG 404
QY 363 GATATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCGCTACTCAACCCCTACTC 422
DB 405 GATATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGAGGCGCTACTCAACCCCTACTC 464
QY 423 GACCCCTACTCAGGCTCGTATGACCCAGAGCTGCCACCATCTCAGCTCCAAATATTC 482
DB 465 GACCCCTACTCAGGCTCGTATGACCCAGAGCTGCCACCATCTCAGCTCCAAATATTC 524
QY 483 CAGCATCTCAGGCTCGTATGACCCAGAGCTGCCACCATCTCAGCTCCAAATATTC 542
DB 525 CAGCATCTCAGGCTCGTATGACCCAGAGCTGCCACCATCTCAGCTCCAAATATTC 584
QY 543 TGTGATGTGAGACAGTGTGCAACAGATTCACACAGTGCACACCCACAGATTCGAT 602
DB 585 TGTGATGTGAGACAGTGTGCAACAGATTCACACAGTGCACACCCACAGATTCGAT 644
QY 603 CAATAGTGAAGGCGGTACACCTGCTCTGACACGACGAGATTTGGCTTCTGGAAAGCCA 662

DB 645 CAATAGTGAAGGCGGTACACCTGCTCTGACACGACGAGATTTGGCTTCTGGAAAGCCA 704
QY 663 GTGCTTAGACATTTGATGATGTGCTGATAGTTACTGACGACGCTCTGTGCGAATTTCC 722
DB 705 GTGCTTAGACATTTGATGATGTGCTGATAGTTACTGACGACGCTCTGTGCGAATTTCC 764
QY 723 TGTATCTATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 782
DB 765 TGTATCTATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
QY 783 CCAAGATGTGACGAGTGTGTCACCGAGAACCCCTGCTGCAACCTGCTCAACCTTA 842
DB 825 CCAAGATGTGACGAGTGTGTCACCGAGAACCCCTGCTGCAACCTGCTCAACCTTA 884
QY 843 CGGCTCTTATCTGCGCCCTGACACCGAGATGATGATGATGATGATGATGATGATGATG 902
DB 885 CGGCTCTTATCTGCGCCCTGACACCGAGATGATGATGATGATGATGATGATGATGATG 944
QY 903 CAGTATATGAGACGAGTGCACCTCTCTGAGTTCCTGCAACATGATGATGATGATGATG 962
DB 945 CAGTATATGAGACGAGTGCACCTCTCTGAGTTCCTGCAACATGATGATGATGATGATG 1004
QY 963 GCCCGCACATATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
DB 1005 GCCCGCACATATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1064
QY 1023 CTGCAAGACATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1082
DB 1065 CTGCAAGACATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
QY 1083 CTCAATTTTACAAAGGCGCTTCAATATGATGATGATGATGATGATGATGATGATGATG 1142
DB 1125 CTCAATTTTACAAAGGCGCTTCAATATGATGATGATGATGATGATGATGATGATGATG 1184
QY 1143 GAGATACATGATTAACCGCTGTATGTCTCTGTAACCTGCTGCAAGACACAGCC 1202
DB 1185 GAGATACATGATTAACCGCTGTATGTCTCTGTAACCTGCTGCAAGACACAGCC 1244
QY 1203 CTTTACCATCTTGTACCGGAGCATGAGAGTGTGTCAGAGAGCTGCTGCTGCTGAT 1262
DB 1245 CTTTACCATCTTGTACCGGAGCATGAGAGTGTGTCAGAGAGCTGCTGCTGCTGAT 1304
QY 1263 CTTCAATATGCAAGGCGACCGCTACCTGCGGCTTATACATTTTCCAGATCAATC 1322
DB 1305 CTTCAATATGCAAGGCGACCGCTACCTGCGGCTTATACATTTTCCAGATCAATC 1364
QY 1323 TGGGATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCCATATAGTCCACCTG 1382
DB 1365 TGGGATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCCATATAGTCCACCTG 1424
QY 1383 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGTGTGAATGATCACTGT 1442
DB 1425 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGTGTGAATGATCACTGT 1484
QY 1443 CAACACTGTCAATCACTTCAAGAGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1502
DB 1485 CAACACTGTCAATCACTTCAAGAGAGCTCCGTGATCCGACTGCGGATATATGTGCGCA 1544
QY 1503 GTACCATTTCTGAGCTCTGCGGCTGAGAGCTCCGAGCGTCCCTCTCAATTTGGCAACAGG 1562
DB 1545 GTACCATTTCTGAGCTCTGCGGCTGAGAGCTCCGAGCGTCCCTCTCAATTTGGCAACAGG 1604
QY 1563 CAGGAGAGAGAGAGAAATACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
DB 1605 CAGGAGAGAGAGAGAAATACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1664
QY 1623 CTGAACGTTTCCCGAAGAGTACGCCGACTTCTGACTCTGACTCTGACTCTGACTCTG 1678
DB 1665 CTGAACGTTTCCCGAAGAGTACGCCGACTTCTGACTCTGACTCTGACTCTGACTCTG 1720

RESULT 11

243 ACAGTGCAGGATGCTTTGACCTGGATGCCGATCAGGACAGTGTGTGATATTATGA 302
 285 ACAGTGCAGGATGCTTTGACCTGGATGCCGATCAGGACAGTGTGTGATATTATGA 344
 303 ATGGCCGACCATCCCGGAGGCTGGCCGAGGACATGATGTGTGTTAAACCAAAATGGCG 362
 345 ATGGCCGACCATCCCGGAGGCTGGCCGAGGACATGATGTGTGTTAAACCAAAATGGCG 404
 363 GATATTATGATTCGCCGACAAACCTGTGTATGAGGAGCCCTACTGCAACCCCTACTC 422
 405 GATATTATGATTCGCCGACAAACCTGTGTATGAGGAGCCCTACTGCAACCCCTACTC 464
 423 GACCCCTACTCAGTCCGCTACCCGACAGCTGCTCCCAACCTGCTGCTCAAACTATCC 482
 465 GACCCCTACTCAGTCCGCTACCCGACAGCTGCTCCCAACCTGCTGCTCAAACTATCC 524
 483 CAGCATCCAGGCTTTATATGCGGCTTTGATACAGATGATGATGATGAAGCAACCAATG 542
 525 CAGCATCCAGGCTTTATATGCGGCTTTGATACAGATGATGATGATGAAGCAACCAATG 584
 543 TGTGATGTGAGCAGTGTGCAACAGATTCACACAGTGCACACCCACAGATTCGAT 602
 585 TGTGATGTGAGCAGTGTGCAACAGATTCACACAGTGCACACCCACAGATTCGAT 644
 603 CAATACGAAAGGCGGTACACCTGCTCCGACCCGACGAGATATGCTTCTGGAAGGCA 662
 645 CAATACGAAAGGCGGTACACCTGCTCCGACCCGACGAGATATGCTTCTGGAAGGCA 704
 663 GTGCTTAGACATTTGATGATGTGCTATGCTATGCTGACGACGCTGTGCGAATGTC 722
 705 GTGCTTAGACATTTGATGATGTGCTATGCTATGCTGACGACGCTGTGCGAATGTC 764
 723 TGTGATCCTATCTTGTATGATGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCTG 782
 765 TGTGATCCTATCTTGTATGATGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCTG 824
 783 CCAGATGTGAAAGCAGTGTGCGACGAGAACCCCTGCTGCAACCTGCTGCAACCTA 842
 825 CCAGATGTGAAAGCAGTGTGCGACGAGAACCCCTGCTGCAACCTGCTGCAACCTA 884
 843 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTTGAGAGATGCGCTTCATG 902
 885 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTTGAGAGATGCGCTTCATG 944
 903 CAGTATATGAGCAGTGTGAGCTTCTGATGCTTCTGCAACATGATGTGGAACA 962
 945 CAGTATATGAGCAGTGTGAGCTTCTGATGCTTCTGCAACATGATGTGGAACA 1004
 1005 GCCCGGACATCTTCTGCTGCGCTCCAGGCTCAAGCTGATCTGCTGATGACCAACG 1064
 1023 CTGCCAAGACATCAACGAATGTGAGCAGAGAACCAACAGTGCACCTGCAAGAGCTG 1082
 1065 CTGCCAAGACATCAACGAATGTGAGCAGAGAACCAACAGTGCACCTGCAAGAGCTG 1124
 1083 CTACAATTTACAAAGGGGCTTCAAAATGATGCAACCCATCGGCTGTGAGAGCTTATCT 1142
 1125 CTACAATTTACAAAGGGGCTTCAAAATGATGCAACCCATCGGCTGTGAGAGCTTATCT 1184
 1143 GAGATTCAGATTAACCGCTGTATGTCTGCTGTGAGAACCTGCTGACAGACAGCC 1202
 1185 GAGATTCAGATTAACCGCTGTATGTCTGCTGTGAGAACCTGCTGACAGACAGCC 1244
 1203 CTTTACCATCTTGTACCGGACATGAGCTGTGTGAGAGAGCTTCCGCTGAGAT 1262
 1245 CTTTACCATCTTGTACCGGACATGAGCTGTGTGAGAGAGCTTCCGCTGAGAT 1304
 1263 CTTTCAAAATGCAAGCAGACCGCTACCTGGGGCTATATTCATTTTCCAGATCAATC 1322
 1305 CTTTCAAAATGCAAGCAGACCGCTACCTGGGGCTATATTCATTTTCCAGATCAATC 1364
 1323 TGGGATGAGGCGAGAAATTTTACATGCGGCAAAAGGGGCCCATCACTGACACCTGCT 1382

1365 TGGGATGAGGCGAGAGATTTTACATGCGGCAAAAGGGCCCATCATGACCTGGT 1424
 1383 GATGACACGCCCATCAAAAGGCCCCGGGAATTCAGCTGAGCTTGAATGATCACTGT 1442
 1425 GATGACACGCCCATCAAAAGGCCCCGGGAATTCAGCTGAGCTTGAATGATCACTGT 1484
 1443 CAACAGTTCATCACTTCAGAGCAGCTCCGTGATCCGACAGCGGATATATGTGCGCA 1502
 1485 CAACAGTTCATCACTTCAGAGCAGCTCCGTGATCCGACAGCGGATATATGTGCGCA 1544
 1503 GTACCCATTTGAGGCTCGGGCTGAGCCCTCCGACGCTCTCTCATTTGGACCAAGGGA 1562
 1545 GTACCCATTTGAGGCTCGGGCTGAGCCCTCCGACGCTCTCTCATTTGGACCAAGGGA 1604
 1563 CAGGAGAAAGAGGAAATTAACAGAGATTAAGAGCAGCAGACAGACAGTATGCAATTC 1622
 1605 CAGGAGAAAGAGGAAATTAACAGAGATTAAGAGCAGCAGACAGACAGTATGCAATTC 1664
 1623 CTGAACGTTTCCCGAGAGAGTACGCCCGGACTTCTGACTCTGACCTGTAATG 1678
 1665 CTGAACGTTTCCCGAGAGAGTACGCCCGGACTTCTGACTCTGACCTGTAATG 1720

RESULT 10
 ABR13627
 ID ABR13627 standard; cDNA; 1720 BP.
 XX AC ABR13627;
 XX AC
 DT 23-APR-2002 (first entry)
 XX DE
 DE Human cDNA encoding extracellular protein-1-like/EGF-like protein, EGF.
 XX
 KM Human, ss; gene; extracellular protein-1-like protein; EGF-like;
 KM protein; epidermal growth factor; EGF; ARCC 97285; gene therapy;
 KM vascular smooth muscle cell proliferation; Marfan syndrome;
 KM wound healing; neurological trauma; acquired immunodeficiency syndrome;
 KM AIDS-related dementia; kidney disorder; liver disorder;
 KM hair follicle growth promotion; burn; ulcer; corneal incision;
 KM corneal inflammation; neoplasm; psoriasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 211..1557
 FT /tag= a
 FT /product= "EGF"
 FT /note= "this CDS minus the stop codon (211-2554) is
 specifically claimed in claim 5"
 FT
 FT sig_peptide 211..285
 FT /tag= b
 FT mat_peptide 286..1554
 FT /tag= c
 FT /label= Mature_EGF
 FT /note= "this region is specifically claimed in
 claim 4"
 XX
 PN US2001051358-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 25-MAR-1999; 9905-0275805.
 XX
 PR 11-APR-1997; 97US-0839525.
 PR 10-APR-1996; 96MO-US05247.
 XX
 PA (OLSEN/) OLSEN H S.
 PA (LIH/) LI H.
 XX
 PI Olsen HS, Li H;
 XX
 DR WPI; 2002-121417/16.

```

Db      1360 TGGGAATGAGGAGAGAAATTTCATGCGGCAAAACGGGCCCAATCAGTGCACCCCTGCT 1419
QY      1383 GATGACACCCCATTAAGAGGCCCCGGGAAATTCAGCTGAGCTTGAATGATCAGCTGT 1442
Db      1420 GATGACACCCCATTAAGAGGCCCCGGGAAATTCAGCTGAGCTTGAATGATCAGCTGT 1479
QY      1443 CAACACTGTCATCACTTCAGAGAGAGCTCGTGTATCCGATCGGAGATATATGTGTGCA 1502
Db      1480 CACACTGTCATCACTTCAGAGAGAGCTCGTGTATCCGATCGGAGATATATGTGTGCA 1539
QY      1503 GTACCCATTCAGAGCTCGGCTGAGAGCTCCGAGAGCTGCTCATATGGACCAAGGGA 1562
Db      1540 GTACCCATTCAGAGCTCGGCTGAGAGCTCCGAGAGCTGCTCATATGGACCAAGGGA 1599
QY      1563 CAGGAGAAAGAGGAATTAACAGAGAAATGAGAGGCAACAGAGCTTATGACATTTCTGT 1622
Db      1600 CAGGAGAAAGAGGAATTAACAGAGAAATGAGAGGCAACAGAGCTTATGACATTTCTGT 1659
QY      1623 CTGAACGTTTCCCGAAGAGTCAAGCCGCACTTCCTACTCTCACTGTACTATTTGACA 1682
Db      1660 CTGAACGTTTCCCGAAGAGTCAAGCCGCACTTCCTACTCTCACTGTACTATTTGACA 1719
QY      1683 CCTGTACCCCTGAGAGAGCTTCCACCCCAAGTTCATATGATACAGTATCAAAAAGTAT 1742
Db      1720 CCTGTACCCCTGAGAGAGTCCACCCCAAGTTCATATGATACAGTATCAAAAAGTAT 1779
QY      1743 ATCATTCGTCCTCGTAGAAGATGTGTGTAATTTTCAAGGCTTCAGTTTATTTTCCA 1802
Db      1780 ATCATTCGTCCTCGTAGAAGATGTGTGTAATTTTCAAGGCTTCAGTTTATTTTCCA 1839
QY      1803 CTATTTTCAAGAAATAGATTAGTGTGCGGGGCTGAGTCAATGTTTCAAGAGCTGTG 1862
Db      1840 CTATTTTCAAGAAATAGATTAGTGTGCGGGGCTGAGTCAATGTTTCAAGAGCTGTG 1899
QY      1863 AACAGCTTGTCTCACTTCTTCACTTCTTCACTTCTTCTTCTCACTGTGTACTGCTTG 1922
Db      1900 AACAGCTTGTCTCACTTCTTCACTTCTTCACTTCTTCTTCTCACTGTGTACTGCTTG 1959
QY      1923 CAAGAGCCCGGGAGCTGGGGGGGAGAACCTGGAGTACTGATTTGCTTTTGGGAGACA 1982
Db      1960 CAAGAGCCCGGGAGCTGGGGGGGAGAACCTGGAGTACTGATTTGCTTTTGGGAGACA 2019
QY      1983 GAGAGGCTATGTAAACAAACACAGAGAGATCGAAGGGTTTTAGAGAAATGTGTTCAA 2042
Db      2020 GAGAGGCTATGTAAACAAACACAGAGAGATCGAAGGGTTTTAGAGAAATGTGTTCAA 2079
QY      2043 AACCATGCTGTGATTTTCAACATATAAGAGTTTCAAGTTGCTCTT 2089
Db      2080 AACCATGCTGTGATTTTCAACATATAAGAGTTTCAAGTTGCTCTT 2126

```

```

RESULT 9
AAV62432
ID      AAV62432 standard; DNA; 1720 BP.
XX      AAV62432;
DT      25-JAN-1999 (first entry)
XX      Human EGF genomic DNA.
XX      Extracellular/epidermal growth factor-like protein; EGF; human; liver;
XX      vascular smooth muscle cell proliferation; neurology; pathology; AIDS;
XX      dementia; ocular; cornea; inflammation; tumor cell; kidney;
XX      wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;
XX      Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasia;
XX      epidermal cell; cancer; psoriasis; detection; ss.
OS      Homo sapiens.
XX      key      Location/Qualifiers
FH      CDS      211..1557

```

```

FT      sig-peptide      /*tag= a
FT      mat-peptide      211..285
FT      /*tag= b
FT      286..1557
FT      /*tag= c
FT      /product= "EGF"
FT      /note= "Extracellular/epidermal growth factor-like"
DB      MO9846746-A1.
DB      22-OCT-1998.
DB      11-APR-1997; 97MO-US06020.
DB      11-APR-1997; 97MO-US06020.
DB      (HUMA-) HUMAN GENOME SCI INC.
DB      LI H, Olsen HS;
DB      WPI: 1998-568728/48.
DB      P-PSDB; AAW79739.
DB      New isolated extracellular/epidermal growth factor - used for
DB      regulating vascular smooth muscle cell proliferation, e.g. for
DB      enhancing neurological functions or treating neoplasia and other
DB      disorders.
DB      Claim 1a: Fig 1A-D; 62pp; English.
DB      This sequence encodes a novel human extracellular/epidermal growth
DB      factor-like protein, EGF. This protein can be used to regulate
DB      vascular smooth muscle cell proliferation and for restoration or
DB      enhancement of neurological functions diminished as a result of other
DB      damaging pathologies such as AIDS dementia. The protein can also be used
DB      to treat senile dementia, ocular disorders such as corneal inflammation,
DB      for targeting tumor cells, for treating kidney disorders, for liver
DB      regeneration or treating liver dysfunction, for treating wounds including
DB      all cutaneous wounds, corneal wounds, and injuries to the
DB      epithelial-lined hollow organs of the body or resulting from trauma such
DB      as burns, abrasions and cuts as well as from surgical procedures such as
DB      surgical incisions and skin grafting. The polypeptides can also be used
DB      for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
DB      other non-healing (trophic) conditions, to treat Marfan syndrome, to
DB      promote hair follicular development, to stimulate growth and
DB      differentiation of various epidermal and epithelial cells in vivo and in
DB      vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
DB      treat neoplasia such as cancers or tumours, skin disorders such as
DB      psoriasis or corneal inflammation. The products can also be used for
DB      identifying EGF receptors, detection, diagnosis and drug screening.
DB      Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other:
DB      Query Match      71.98; Score 1674.2; DB 19; Length 1720;
DB      Best Local Similarity 99.88; Pred. No. 0;
DB      Matches 1673; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      3 CCCGGGCTCTCCCGCTGCTCTCCAGAGATCGCTGCGGCCCTGTGAATTAACACCC 62
Db      45 CCCGGGCTCTCCCGCTGCTCTCCAGAGATCGCTGCGGCCCTGTGAATTAACACCC 104
QY      63 GCGAGCCCGGAGGCGCAGAGAGAGGCGAGCTGCGGCCGAGCTCTCGGGGGTCCGCCCG 122
Db      105 GCGAGCCCGGAGGCGCAGAGAGAGGCGAGCTGCGGCCGAGCTCTCGGGGGTCCGCCCG 164
QY      123 CGAGCTTCTTCTCGCTTGGCATCTCTCTCGCGGCTCTTGGCATCTCCAGAAATAA 182
Db      165 CGAGCTTCTTCTCGCTTGGCATCTCTCTCGCGGCTCTTGGCATCTCCAGAAATAA 224
QY      183 AAGGATACACAGTATTCATATTCGGGCTCTGCTCTCCAGAGCCCTGGGAATGCACAGC 242
Db      225 AAGGATACACAGTATTCATATTCGGGCTCTGCTCTCCAGAGCCCTGGGAATGCACAGC 284

```


404 CAGCTTTCTTCCGCTTCGATCTCTCTCGCGGCTTGGACATGCCAGAAATAA 463
183 AAGGATACACTGTTACATTCGCTCTCTCTTCCAAAGCCCTGGGAATGCACAGC 242
464 AAGGATACACTGTTACATTCGCTCTCTCTTCCAAAGCCCTGGGAATGCACAGC 523
243 ACAGTGCAGGAATGGCTTGGACCTGATCGCCAGTGCAGGAGTGTATGATGATGA 302
524 ACAGTGCAGGAATGGCTTGGACCTGATCGCCAGTGCAGGAGTGTATGATGATGA 583
303 ATGCCCAACCATCCCGAGGCGCTGCCAGAGACATGATGTGTATTAACCAAAATGGCG 362
584 ATGCCCAACCATCCCGAGGCGCTGCCAGAGACATGATGTGTATTAACCAAAATGGCG 643
363 GATTTATGACATTCGCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTGC 422
644 GATTTATGACATTCGCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGGAACCCCTACTGC 703
423 GACCCCTACTAGTCCGATCCGATCCAGAGAGCTGCCCAACACTCTCAGCTCCAAACTATCC 482
704 GACCCCTACTAGTCCGATCCGATCCAGAGAGCTGCCCAACACTCTCAGCTCCAAACTATCC 763
483 CAGAGTCTCAGGCTCTTATATGCGCTTGGATACCAATGATGAAGAACCAATG 542
764 CAGAGTCTCAGGCTCTTATATGCGCTTGGATACCAATGATGAAGAACCAATG 823
543 TGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCCAGATCTGCAT 602
824 TGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCCAGATCTGCAT 883
603 CAATCTGAAGCGGGGTACACTGCTCTGACACCGAGAGATTTGGCTTGTGAAGGCCA 662
884 CAATCTGAAGCGGGGTACACTGCTCTGACACCGAGAGATTTGGCTTGTGAAGGCCA 943
663 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTCTGTGGAAATGTTCC 722
944 GTGCTAGACATTTGATGAATGTGCTATGTTACTGACAGAGCTCTGTGGAAATGTTCC 1003
723 TGGATCTCTATTTCTTGTACATGCAACCTGGTTTACCCTAATGAGATGGAAGTCTTG 782
1004 TGGATCTCTATTTCTTGTACATGCAACCTGGTTTACCCTAATGAGATGGAAGTCTTG 1063
783 CCAAGATGTGAAGAGATGTGCCAACCGAAGACCCCTGGCTGCAACCTGCTCAACCTGA 842
1064 CCAAGATGTGAAGAGATGTGCCAACCGAAGACCCCTGGCTGCAACCTGCTCAACCTGA 1123
843 CGGCTCTTCATCTGCGGCTGTACCCAGATATGATGAGAAATGGCGTTCAATG 902
1124 CGGCTCTTCATCTGCGGCTGTACCCAGATATGATGAGAAATGGCGTTCAATG 1183
903 CAGTATATGAGAGATGTGCAAGTCTTGTGATCTCTGCAACATGATGTGTGAACCA 962
1184 CAGTATATGAGAGATGTGCAAGTCTTGTGATCTCTGCAACATGATGTGTGAACCA 1243
963 GCCCGGACATATCTTGTCTCTGCTCCCTCAGGCTACATCTGCTGTGATGAACCGAAG 1022
1244 GCCCGGACATATCTTGTCTCTGCTCCCTCAGGCTACATCTGCTGTGATGAACCGAAG 1303
1023 CTGCCAAGACATCAAGAAATGTAGCAGAGAAACACAGTGAACCGTGCAGAGAGCTG 1082
1304 CTGCCAAGACATCAAGAAATGTAGCAGAGAAACACAGTGAACCGTGCAGAGAGCTG 1363
1083 CTACAAATTTACAAGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGGCTTATCT 1142
1364 CTACAAATTTACAAGGGGCTTCAAAATGCATGACCCCATCCGCTGTGAGAGGCTTATCT 1423
1143 GAGGATCATGATTAACCGCTGTATGTCTGCTGAGAAACCCCTGGCTGCAGAGACAGCC 1202
1424 GAGGATCATGATTAACCGCTGTATGTCTGCTGAGAAACCCCTGGCTGCAGAGACAGCC 1483
1203 CTTTACCATCTTGTACCGGACATGAGAGTGTGTGAGAGAGCTCCGTTCCCGGTGACAT 1262
1484 CTTTACCATCTTGTACCGGACATGAGAGTGTGTGAGAGAGCTCCGTTCCCGGTGACAT 1543

1263 CTTCAAAATGCAAGCCACGACCCGCTACCCCTGGGCGCTATTAATTTCCAGATCAAAATC 1322
1544 CTTCAAAATGCAAGCCACGACCCGCTACCCCTGGGCGCTATTAATTTCCAGATCAAAATC 1603
1323 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGT 1382
1604 TGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGT 1663
1383 GATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTTGGAAATGATACCTGT 1442
1664 GATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGACCTTGGAAATGATACCTGT 1723
1443 CAACACTGTATCAACTTCAGAGGAGCTCCGTGATCCGAGTGGGGAATATGTCGCA 1502
1724 CAACACTGTATCAACTTCAGAGGAGCTCCGTGATCCGAGTGGGGAATATGTCGCA 1783
1503 GTACCCATTTCTGAGGCTCGGGCTGGAGCTCCGAGCTGCTCTCATTTGGCACAAAGGA 1562
1784 GTACCCATTTCTGAGGCT----- 1800
1563 CAGGAGAGAGAGAAATTAACAGAGAAATGAGAGCCGACAGACGTTAAGCATTTCTG 1622
1801 -----CTG 1803
1623 CTGAAGCTTTCCCGAAGAGTCAAGCCCGACTCTCTGACTCTACCTGTACTATTGACA 1682
1804 CTGAAGCTTTCCCGAAGAGTCAAGCCCGACTCTCTGACTCTACCTGTACTATTGACA 1863
1683 CTTGTACACCTGCAAGACTTGCACACCCCACTTCTATGATTAACATTAACAAAGATTT 1742
1864 CTTGTACACCTGCAAGACTTGCACACCCCACTTCTATGATTAACATTAACAAAGATTT 1923
1743 ATCATGTCCCGCTGATGAGAAATGTTGGAATTTCAAGGCTTACGTTATTTCA 1802
1924 ATCATGTCCCGCTGATGAGAAATGTTGGAATTTCAAGGCTTACGTTATTTCA 1983
1803 CTAATTTCAAGAAATATAGATTAGTTTGCAGGGGTGAGACTCTATGTTCAAAAGCTGTG 1862
1984 CTAATTTCAAGAAATATAGATTAGTTTGCAGGGGTGAGACTCTATGTTCAAAAGCTGTG 2043
1863 AACAGCTGTCTGACCTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1922
2044 AACAGCTGTCTGACCTCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2103
1923 CAAGAGCCCGGAGCTGGCGGGGAACCTGGGAGTACGTACTTTTGGCTTGGCTGACCA 1982
2104 CAAGAGCCCGGAGCTGGCGGGGAACCTGGGAGTACGTACTTTTGGCTTGGCTGACCA 2163
1983 GAGAAAGCTATGTAACAAACACAGAGATGCAAGGCTTTTGAAGATGTGTTCA 2042
2164 GAGAAAGCTATGTAACAAACACAGAGATGCAAGGCTTTTGAAGATGTGTTCA 2223
2043 AACCATGCTGTATTTTCAACCATTAAGAAAGTTCAAGTGTCTTAAATTTGTATAC 2102
2224 AACCATGCTGTATTTTCAACCATTAAGAAAGTTCAAGTGTCTTAAATTTGTATAC 2283
2103 GGTTAATCTGTCTGTTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 2162
2284 GGTTAATCTGTCTGTTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 2343
2163 GAAAGGCTTCAGACATGCTATGTTCTGCTTCCAAACCCAGTCTCTCCATTTT 2222
2344 GAAAGGCTTCAGACATGCTATGTTCTGCTTCCAAACCCAGTCTCTCCATTTT 2403
2223 AGCCGAGTGTCTTTTGAAGAGCCCTTAATCTTGTCTTCTTGAATTTTGAATTTTGA 2282
2404 AGCCGAGTGTCTTTTGAAGAGCCCTTAATCTTGTCTTCTTGAATTTTGAATTTTGA 2463
2283 GATTTGGAATGAGAGGCTTCAAACTGATTAATTTTGAAGGA 2328
2464 GATTTGGAATGAGAGGCTTCAAACTGATTAATTTTGAAGGA 2509

PI Wood WI, Zhang Z;
XX WPI; 2001-050091/06.
DR P-PSDB; AAB31183.
XX
PT Isolated nucleic acid molecule encoding a PRO polypeptide which is a
PT transmembrane polypeptide is useful for gene therapy and identification
PT of related polypeptides -
XX
XX
PS Claim 2; Fig 9; 244pp; English.
CC The present sequence encodes a human secreted and transmembrane
CC polypeptide. The specification describes human polypeptides, designated
CC PRO156, PRO444, PRO183, PRO185, PRO210, PRO215, PRO217, PRO242, PRO288,
CC PRO355, PRO361, PRO308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
CC PRO3248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
CC PRO337, PRO411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells
CC can be modulated with agents that bind to these polypeptides, resulting
CC in the death of the cells. The polynucleotides encoding these
CC polypeptides are useful in the recombinant production of the
CC polypeptides, as a hybridisation probe to screen libraries to isolate
CC homologous sequences, or to map the gene. They may also be used for
CC analysing genetic disorders, and to produce transgenic animals which are
CC useful for the development and screening of therapeutically useful
CC reagents. The polynucleotides can also be used in gene therapy e.g. to
CC replace a defective gene.
XX
SQ Sequence 2609 BP; 622 A; 718 C; 622 G; 647 T; 0 other;
Query Match 99.88; Score 2322.8; DB 22; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
3 CCGGGGCTTCCTCCGCTCTCCACGACTCGCTGCGCCCTTGGAATAAACACC 62
284 CCGGGGCTTCCTCCGCTCTCCACGACTCGCTGCGCCCTTGGAATAAACACC 343
63 GCGAGCCCCGAGGCCCCAGAGAGAGCGAGCTGCCCCAGCTCTCCGGGGTCCGCCCG 122
344 GCGAGCCCCGAGGCCCCAGAGAGAGCGAGCTGCCCCAGCTCTCCGGGGTCCGCCCG 403
123 CGAGCTTCTTCTGCGCTTGCGATCTCTCTCGCGCTCTTGAGCATGCCAGATAAA 182
404 CGAGCTTCTTCTGCGCTTGCGATCTCTCTCGCGCTCTTGAGCATGCCAGATAAA 463
183 AAGGATACGACTGTTACATCTGCGCTCTGCTCTTCCAAAGCCCTGGGAATGACAGGC 242
464 AAGGATACGACTGTTACATCTGCGCTCTGCTCTTCCAAAGCCCTGGGAATGACAGGC 523
243 ACAGTCACGAAATGCTTTGACCTGATCGCCAGTCAGACAGAGTGTATAGATTGATGA 302
524 ACAGTCACGAAATGCTTTGACCTGATCGCCAGTCAGACAGAGTGTATAGATTGATGA 583
303 ATGCCGACCATCCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 362
584 ATGCCGACCATCCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 643
363 GTATTATGATTCCTCCGGAACAAACCTGTATCGAGGCTTACTGAAACCTTACTGC 422
644 GTATTATGATTCCTCCGGAACAAACCTGTATCGAGGCTTACTGAAACCTTACTGC 703
423 GACCCCTACTAGTCCGATACCGAGAGCTGCCCACTCTCAGCTCCAAATCATCC 482
704 GACCCCTACTAGTCCGATACCGAGAGCTGCCCACTCTCAGCTCCAAATCATCC 763
483 CACGATCTCAGGCTCTTATATGCGCTTTGATACCATGATGATGAAGCAACCAATG 542
764 CACGATCTCAGGCTCTTATATGCGCTTTGATACCATGATGATGAAGCAACCAATG 823
543 TGTGATGTGAGAGAGTGTACAGATGCCACAGTCCACAGTCCACCCAGATCTGCAT 602
824 TGTGATGTGAGAGAGTGTACAGATGCCACAGTCCACAGTCCACCCAGATCTGCAT 883

QY 603 CAATACGTAAGGCGGGTACACCTGCTCTGCACCGAGGATATGCTTGTGAGAGCCA 662
DB 884 CAATACGTAAGGCGGGTACACCTGCTCTGCACCGAGGATATGCTTGTGAGAGCCA 943
QY 663 GTGCTTACATGATGATGATGCTGCTATGCTTACTGCGACGACCTGTGCGAATGTC 722
DB 944 GTGCTTACATGATGATGATGCTGCTATGCTTACTGCGACGACCTGTGCGAATGTC 1003
QY 723 TGGATCCATCTCTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 782
DB 1004 TGGATCCATCTCTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 1063
QY 783 CCAAGATGTGAACGAGTGTGCGACGAGAGAACCTGCGTCAACACCTA 842
DB 1064 CCAAGATGTGAACGAGTGTGCGACGAGAGAACCTGCGTCAACACCTA 1123
QY 843 CGGCTCTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAGAGATGCGCTTCAATG 902
DB 1124 CGGCTCTTTCATCTGCGCTGTGACCCAGAGATGAACTTGAGAGATGCGCTTCAATG 1183
QY 903 CAGTATATGAGACGAGTGCAGTCTGCTGATGCTCTGCTCAACATGAGTGTGACCA 962
DB 1184 CAGTATATGAGACGAGTGCAGTCTGCTGATGCTCTGCTCAACATGAGTGTGACCA 1243
QY 963 GCCCGGACATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
DB 1244 GCCCGGACATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1303
QY 1023 CTGCAAGACATCAACGATGTGACGACAGAACACACGTCGACGACGACGACGACG 1082
DB 1304 CTGCAAGACATCAACGATGTGACGACAGAACACACGTCGACGACGACGACGACG 1363
QY 1083 CTCAATTTAAAGGGGCTTCAATGATGACACCCATCCCTGTGAGAGGCTTATCT 1142
DB 1364 CTCAATTTAAAGGGGCTTCAATGATGACACCCATCCCTGTGAGAGGCTTATCT 1423
QY 1143 GAGATGATGATTAACCCCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202
DB 1424 GAGATGATGATTAACCCCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1483
QY 1203 CTTTACCATCTTGTACCGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262
DB 1484 CTTTACCATCTTGTACCGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1543
QY 1263 CTTTACCATCTTGTACCGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1322
DB 1544 CTTTACCATCTTGTACCGGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1603
QY 1323 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATGATGCGACCTGT 1382
DB 1604 TGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGCGCCCATGATGCGACCTGT 1663
QY 1383 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGCTTGAATATGATCACTGT 1442
DB 1664 GATGACAGCGCCCATCAAAAGGCGCCGGAATTCAGCTGAGCTTGAATATGATCACTGT 1723
QY 1443 CAACACGTGTATCAATCTCAAGAGCAGCTCGGTATCCGACCTGCGGATATATGTGCGCA 1502
DB 1724 CAACACGTGTATCAATCTCAAGAGCAGCTCGGTATCCGACCTGCGGATATATGTGCGCA 1783
QY 1503 GTACCCATCTGAGACCTGCGGCTGAGGCTCGGAGCTCGGAGCTCTCATTTGGACCAAGGA 1562
DB 1784 GTACCCATCTGAGACCTGCGGCTGAGGCTCGGAGCTCGGAGCTCTCATTTGGACCAAGGA 1843
QY 1563 CAGGAGAGAGAGAGAAATTAACAGAGATGAGACGACACAGCTTATGAGCATTTCTGT 1622
DB 1844 CAGGAGAGAGAGAGAAATTAACAGAGATGAGACGACACAGCTTATGAGCATTTCTGT 1903
QY 1623 CTGAAGCTTTCCTCGGAGAGAGCAGCCCGGACTTCTGACTCTGACCTGTACATATGCGCA 1682
DB 1904 CTGAAGCTTTCCTCGGAGAGAGCAGCCCGGACTTCTGACTCTGACCTGTACATATGCGCA 1963

```

1083 CTACAAATTTACAAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTATCT 1142
1364 CTACAAATTTCAAGGGGGCTTCAATATGATGACCCCATCCGCTGTGAGAGACCTATCT 1423
1143 GAGGATGATGATTAACCGCTGTATGTCTCTGTGAGAAACCTGTGCTGACAGACAGCC 1202
1424 GAGGATGATGATTAACCGCTGTATGTCTCTGTGAGAAACCTGTGCTGACAGACAGCC 1483
1203 CTTTACATCTTGTACCGGGGACATGACGATGTGTGACAGAGCTTCCTCCGCTGACAT 1262
1484 CTTTACATCTTGTGTACCGGGGACATGACGATGTGTGACAGAGCTTCCTCCGCTGACAT 1543
1263 CTTTCAATATGCAAGCCAGACCCGCTACCCCTGGGGCCATTACATTTTCCAGATCAATC 1322
1544 CTTTCAATATGCAAGCCAGACCCGCTACCCCTGGGGCCATTACATTTTCCAGATCAATC 1603
1323 TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGCCCATCACTGCGCCTGTG 1382
1604 TGGGAATGAGGCGACAGAAATTTTACATGCGGCAAAAGGGCCCATCACTGCGCCTGTG 1663
1383 GATGACACGGCCCATCAAGGGGGCGGGAATTCAGCTGTGACCTTGGAAATGATACATCT 1442
1664 GATGACACGGCCCATCAAGGGGGCGGGAATTCAGCTGTGACCTTGGAAATGATACATCT 1723
1443 CAACACTGTATCAACTTTCAGAGGACAGCTCGTGTATCCGACTGCGGATATATGTGCGA 1502
1724 CAACACTGTATCAACTTTCAGAGGACAGCTCGTGTATCCGACTGCGGATATATGTGCGA 1783
1503 GTACCAATCTGTAGCCCTCGGGCTGTGAGACCTTCGACGCTCTTCATTTGGACCAAGGA 1562
1784 GTACCAATCTGTAGCCCTCGGGCTGTGAGACCTTCGACGCTCTTCATTTGGACCAAGGA 1843
1563 CAGGAGAGAGAGAAATACAGAGAGATGAGAGCGACAGACGTTAGGATTCCTG 1622
1844 CAGGAGAGAGAGAAATACAGAGAGATGAGAGCGACAGACGTTAGGATTCCTG 1903
1623 CTGAAGCTTCCCGGAGAGAGTCAAGCCGACTTCTCTACTCTCACTGTACTATTTGAGA 1682
1904 CTGAAGCTTCCCGGAGAGAGTCAAGCCGACTTCTCTACTCTCACTGTACTATTTGAGA 1963
1683 CCTGTACCCCTGTGAGAGACTTGCACCCGACTTCTATGATACGATTATCAAAAAGTAT 1742
1964 CCTGTACCCCTGTGAGAGACTTGCACCCGACTTCTATGATACGATTATCAAAAAGTAT 2023
1743 ATCATTTGCTCCCTGTATAGAAATGTGTGGAATTTTCAAGGCTTCAGTTTATTTTCA 1802
2024 ATCATTTGCTCCCTGTATAGAAATGTGTGGAATTTTCAAGGCTTCAGTTTATTTTCA 2083
1803 CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTCTATGTTTCAAGAGCTGTG 1862
2084 CTATTTTCAAGAAATAGATTAAGTTTGGGGGGTCTGAGTCTATGTTTCAAGAGCTGTG 2143
1863 AACACTTGTCTGTACATTTCTTCACTCTTCTCTCACTGTCTGTCTGTACTGTG 1922
2144 AACACTTGTCTGTACATTTCTTCACTCTTCTCTCACTGTCTGTCTGTACTGTG 2203
1923 CAAGAAGCCGGAGAGTGGGGGGAGACCTGGAGTACTGTTGCTTTTGGCGTACGA 1982
2204 CAAGAAGCCGGAGAGTGGGGGGAGACCTGGAGTACTGTTGCTTTTGGCGTACGA 2263
1983 GAGAAAGGCTATGTAAACAAACACAGCAGAGATGAAAGGTTTGTAGAAATGTGTTTCA 2042
2264 GAGAAAGGCTATGTAAACAAACACAGCAGAGATGAAAGGTTTGTAGAAATGTGTTTCA 2323
2043 AACCATGCTGTATTTTCAACCATTAAGAGAGTTGCTGCTTAAATTTGTATTAAC 2102
2324 AACCATGCTGTATTTTCAACCATTAAGAGAGTTGCTGCTTAAATTTGTATTAAC 2383
2103 GGTAAATCTGCTGTATTTTCAATTTTGTATTTTAAATAATNGTGTAGAAATCTCTC 2162
2384 GGTAAATCTGCTGTATTTTCAATTTTGTATTTTAAATAATNGTGTAGAAATCTCTC 2443
2163 GAAAGGCTTTCAGACACATGCTATGTTTCTGTCTTCCCAACCAAGTCTCCTCCATTTT 2222
```

```

Db 2444 GAAAGGCTTTCAGACACATGCTATGTTCTGTCTTCCCAACCAAGTCTCCTCCATTTT 2503
QY 2223 AGCCCACTGTTTCTTTGAGACCCCTTAATCTTGTCTTCTTTAGAAATTTTACCAATT 2282
Db 2504 AGCCCACTGTTTCTTTGAGACCCCTTAATCTTGTCTTCTTTAGAAATTTTACCAATT 2563
QY 2283 GATTTGAAATGACAGAGTCTTCAACATGATTAATTTTGAAGA 2328
Db 2564 GATTTGAAATGACAGAGTCTTCAACATGATTAATTTTGAAGA 2609

RESULT 6
AAC86968
ID AAC86968 standard; cdna; 2609 BP.
XX AAC86968;
AC AAC86968;
XX AAC86968;
DT 20-APR-2001 (first entry)
XX 20-APR-2001 (first entry)
DE Nucleotide sequence of human polypeptide PRO210.
XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
XX PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1308;
XX PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
XX PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
XX PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
XX PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;
XX ss.
XX Homo sapiens.
XX
FH Key 450..1796
FT CDS /*tag= a
FT sig_peptide 450..524
FT /*tag= b
PN MO200077037-A2.
XX
PD 21-DEC-2000.
PF 22-MAY-2000; 2000MO-US14042.
XX
PR 15-JUN-1999; 99US-0139695.
PR 20-JUL-1999; 99US-0145070.
PR 26-JUL-1999; 99US-0145698.
PR 17-AUG-1999; 99US-0149396.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28301.
PR 02-DEC-1999; 99MO-US28565.
PR 07-DEC-1999; 99US-0169495.
PR 05-JAN-2000; 2000MO-US00219.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05841.
PR 20-MAR-2000; 2000MO-US07377.
PR 30-MAR-2000; 2000MO-US08439.
PR 15-MAY-2000; 2000MO-US13358.
PR 17-MAY-2000; 2000MO-US13705.
XX
PA (GENE) GENENTECH INC.
XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Gurney AL, Kijavrin IV, Mather JP, Napier MA, Pan J;
PI Paoul NF, Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM;
```

PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000MO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15284.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000MO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000MO-US23328.
 PR 08-NOV-2000; 2000MO-US30952.
 PR 01-DEC-2000; 2000MO-US32678.
 PR 20-DEC-2000; 2000MO-US34956.
 (GETH) GENENTECH INC.
 Baker KP, Chen J, Desnoyers L, Goddard A, Godowski P, Gurney AL;
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 MPI: 2001-602746/68.
 P-PSDB; AA029227.
 Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 presence of tumours, such as prostate and breast tumours, in mammals and
 to screen for modulators of the compounds -
 Claim 2: Fig 407; 774pp; English.
 Sequences AAS45925-AA846231 represent DNA molecules encoding and PCR
 primers for PRO polypeptides of the invention. The sequences of the
 invention can be used to detect the presence of a tumour in a mammal by
 comparing the level of expression of a PRO polypeptide in a test sample
 of cells from the animal and a control sample of normal cells, whereby a
 higher level of expression in the test sample indicates the presence of a
 tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
 pigs, goats and rabbits but are preferably human. The polypeptides can be
 used to stimulate tumour necrosis factor (TNF) alpha release from human
 blood, when contacted with it. A specific polypeptide can be used to
 stimulate the proliferation or differentiation of chondrocyte cells. The
 PRO proteins can be used to determine the presence of tumours and also
 susceptibility to tumour development, particularly adrenal, lung, colon,
 breast, prostate, rectal, cervical, or liver tumours, in mammalian
 subjects. The oligonucleotide probes specific for the PRO nucleic acids
 can be used for genetic analysis of individuals with genetic disorders.
 Sequence 2609 BP; 622 A; 718 C; 622 G; 647 T; 0 other;

Query Match 99.8%; Score 2322.8; DB 22; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	3	CCGCGCGCTCTCCCGCTCTCTCTCAAGACATCGCTGGCCCTCTGGAATAAACACCC	62
DB	284	CCCGGCGCTCTCCCGCTCTCTCTCTCAAGACATCGCTGGCCCTCTGGAATAAACACCC	343
QY	63	GGCAGCCCCCAGAGGCCCCAGAGAGGCGCAGCTGCCGAGCTCTCCGGGGGTCCGCCCG	122
DB	344	GGCAGCCCCCAGAGGCCCCAGAGAGGCGCAGCTGCCGAGCTCTCCGGGGGTCCGCCCG	403
QY	123	CGAGCTTCTTCTCGCCCTGCGATCTCTCTCGCGGCTTGGACATCCGGAATAA	182
DB	404	CGAGCTTCTTCTCGCCCTGCGATCTCTCTCGCGGCTTGGACATCCGGAATAA	463
QY	183	AAGGATACCTACTTACATCTTGGCTCTGCTCTTCCAAAGCCCTGGGAATGACAGGC	242
DB	464	AAGGATACCTACTTACATCTTGGCTCTGCTCTTCCAAAGCCCTGGGAATGACAGGC	523
QY	243	ACAGTGCAGAGAGGCTTTGACCTGGATGCGCAGTCCAGACAGTGTATGATGATGA	302
DB	524	ACAGTGCAGAGAGGCTTTGACCTGGATGCGCAGTCCAGACAGTGTATGATGATGA	583
QY	303	ATGCCGAACCATCCCGGCGCTGCGGAGAGACATGATGTGTAAACCAATGAGCG	362
DB	584	ATGCCGAACCATCCCGGCGCTGCGGAGAGACATGATGTGTAAACCAATGAGCG	643
QY	363	GTAATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTCTACTCGAACCCCTACTC	422
DB	644	GTAATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTCTACTCGAACCCCTACTC	703
QY	423	GACCCCTTACTAGGTCGCTGACCCAGCAGCTGCCCCACACTCTCAAGCTCAACATATGC	482
DB	704	GACCCCTTACTAGGTCGCTGACCCAGCAGCTGCCCCACACTCTCAAGCTCAACATATGC	763
QY	483	CACGATCTCCAGGCTCTTATATGCGGCTTGGATGATGATGATGATGATGATGATG	542
DB	764	CACGATCTCCAGGCTCTTATATGCGGCTTGGATGATGATGATGATGATGATGATG	823
QY	543	TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	602
DB	824	TGTGATGTGAGAGAGTGTGCAACAGATTCACACAGTGCACACCCACAGATCTGCAT	883
QY	603	CAATCTGGAAGGCGGCTGACCTGCTCTGACACGAGATTTGGCTCTTGGAGGCCA	662
DB	884	CAATCTGGAAGGCGGCTGACCTGCTCTGACACGAGATTTGGCTCTTGGAGGCCA	943
QY	663	GTGCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	722
DB	944	GTGCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1003
QY	723	TGATTCCTATTCCTGTACATGCAACCCGTTTACCCCTCAATGAGATGAGATGATG	782
DB	1004	TGATTCCTATTCCTGTACATGCAACCCGTTTACCCCTCAATGAGATGAGATGATG	1063
QY	783	CCAAGATGTGAGAGAGTGTGCAACCCGCTGCTGCAACCTGCTGCAACCTGCA	842
DB	1064	CCAAGATGTGAGAGAGTGTGCAACCCGCTGCTGCAACCTGCTGCAACCTGCA	1123
QY	843	CGGCTCTTATGATGCGGCTGACCCGAGATGATGATGATGATGATGATGATG	902
DB	1124	CGGCTCTTATGATGCGGCTGACCCGAGATGATGATGATGATGATGATGATG	1183
QY	903	CAGTATGATGAGAGAGTGTGCAACCCGCTGCTGCAACCTGCTGCAACCTGCA	962
DB	1184	CAGTATGATGAGAGAGTGTGCAACCCGCTGCTGCAACCTGCTGCAACCTGCA	1243
QY	963	GGCCGGGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1022
DB	1244	GGCCGGGACATCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1303
QY	1023	CTGCAAGACATCAAGCAATGTGAGACAGAGAACCAACAGTGCACAGCTGACAGAG	1082
DB	1304	CTGCAAGACATCAAGCAATGTGAGACAGAGAACCAACAGTGCACAGCTGACAGAG	1363

D 756 TGTGATGTGGACGAGTGTGCAACGATTTCCACAGTGCACCCACGATCCACCGATCTGCAT 815
Q 603 CAATACATGAAAGCGGGTACACCTGCTCTGTGACCGAGGATATGCTTCTGGAAGCCA 662
D 816 CAATACATGAAAGCGGGTACACCTGCTCTGTGACCGAGGATATGCTTCTGGAAGCCA 875
Q 663 GTGCTTACACATGATGAAATGTCGCTATGCTTACGCGCAGAGCTCTGCGGAATGCTTC 722
D 876 GTGCTTACACATGATGAAATGTCGCTATGCTTACGCGCAGAGCTCTGCGGAATGCTTC 935
Q 723 TGAATCCATATCTGTACATGCAACCTGCTTATACCTCAATGAGAGATGGAAGCTTGG 782
D 936 TGGATCTATCTGTGTACATGCAACCTGCTTATACCTCAATGAGAGATGGAAGCTTGG 995
Q 783 CCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACTTA 842
D 996 CCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACTTA 1055
Q 843 CGGCTCTTCAATCTGCGCCCTGTGACCGAGATATGAACTTGAGGAAGATGGCTTCATGG 902
D 1056 CGGCTCTTCAATCTGCGCCCTGTGACCGAGATATGAACTTGAGGAAGATGGCTTCATGG 1115
Q 903 CAGTGAATGTGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
D 1116 CAGTGAATGTGACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1175
Q 963 GCCCGGACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
D 1176 GCCCGGACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1235
Q 1023 CTGCAACATGCAACGATGAGACAGAGAACCAAGTGCACCTGAGAGAGAGAGTGTG 1082
D 1236 CTGCAACATGCAACGATGAGACAGAGAACCAAGTGCACCTGAGAGAGAGAGTGTG 1235
Q 1083 CTACAAATTTAAGAGGGGCTTCAAAATGATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCT 1142
D 1296 CTACAAATTTAAGAGGGGCTTCAAAATGATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCT 1355
Q 1143 GAGATCAATGATTAACCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
D 1356 GAGATCAATGATTAACCGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
Q 1203 CTTCACATCTTGTACCGGGAATGAGAGAGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
D 1416 CTTCACATCTTGTACCGGGAATGAGAGAGTGTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1475
Q 1263 CTTCACAAATGCAAGCGACGACCCGCTACCCCTGCGGCTTATTTACATTTTCCAGATCAAAATC 1322
D 1476 CTTCACAAATGCAAGCGACGACCCGCTACCCCTGCGGCTTATTTACATTTTCCAGATCAAAATC 1535
Q 1323 TGGGAATGAGGCGAAGAAATTTTACATGAGGGAAGAGGGGCTCATGTCGCCACCTGGT 1382
D 1536 TGGGAATGAGGCGAAGAAATTTTACATGAGGGAAGAGGGGCTCATGTCGCCACCTGGT 1595
Q 1383 GATGACAGCGCCCATCAAAAGGCGCCGGAATCAAGCTGAGCTTGGAAATGATCACTGT 1442
D 1596 GATGACAGCGCCCATCAAAAGGCGCCGGAATCAAGCTGAGCTTGGAAATGATCACTGT 1655
Q 1443 CAACATCTGATCAACTTCAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1502
D 1656 CAACATCTGATCAACTTCAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1715
Q 1503 GTACCATTTCTGAGCTGCGGCTGAGCTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1562
D 1716 GTACCATTTCTGAGCTGCGGCTGAGCTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1775
Q 1563 CAGGAGAGAGAGAAATTAACAGAGAGATGAGAGGACACAGAGCTTAGGCAATTTCTCTG 1622
D 1776 CAGGAGAGAGAGAAATTAACAGAGAGATGAGAGGACACAGAGCTTAGGCAATTTCTCTG 1835
Q 1623 CTGAACGTTTCCCGAAGAGTCAAGCCGCACTTCTGAGCTCTCACTGTAATTTGACGA 1682
D 1836 CTGAACGTTTCCCGAAGAGTCAAGCCGCACTTCTGAGCTCTCACTGTAATTTGACGA 1895

Q 1683 CCTGTACCCCTGACAGACTTGGCAACCCAGTTCATATGATACAGTTATCAAAAGTATT 1742
D 1896 CCGTCACCCCTGACAGACTTGGCAACCCAGTTCATATGATACAGTTATCAAAAGTATT 1955
Q 1743 ATCATTTGCTCCCTGATAGAAATTTGTTGTAATTTTCAAGGCTTCACTTATTTTCA 1802
D 1956 ATCATTTGCTCCCTGATAGAAATTTGTTGTAATTTTCAAGGCTTCACTTATTTTCA 2015
Q 1803 CTATTTTCAAAAGAAATATAGTTAGTTGGGGGGCTGTGAGTCTATGTTCAAGACTGTG 1862
D 2016 CTATTTTCAAAAGAAATATAGTTAGTTGGGGGGCTGTGAGTCTATGTTCAAGACTGTG 2075
Q 1863 AACAGCT 1922
D 2076 AACAGCT 2135
Q 1923 CAAGACCCGGGAGCTGCGGGGAAACCTGGGAGTGTGAGTGTGCTTTTGGCTAGACA 1982
D 2136 CAAGACCCGGGAGCTGCGGGGAAACCTGGGAGTGTGAGTGTGCTTTTGGCTAGACA 2195
Q 1983 GAGAGGCTATGTAAACAAACACAGAGAGTGAAGGGTTTGTAGAAATGTGTCA 2042
D 2196 GAGAGGCTATGTAAACAAACACAGAGAGTGAAGGGTTTGTAGAAATGTGTCA 2255
Q 2043 AACCATGCT 2102
D 2256 AACCATGCT 2315
Q 2103 GGTAAATCTGCT 2162
D 2316 GGTAAATCTGCT 2375
Q 2163 GAAAGGCTTTCAGACATGCTATGTTCTGCTTCCCAACCCAGTCTCCCTCCATTTT 2222
D 2376 GAAAGGCTTTCAGACATGCTATGTTCTGCTTCCCAACCCAGTCTCCCTCCATTTT 2435
Q 2223 AGCCAGTGTGCTTGTGAGGACCCCTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2282
D 2436 AGCCAGTGTGCTTGTGAGGACCCCTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2495
Q 2283 GGATTGGAATGCAAGAGTCTCAAACTGATTAATATTGGAAGAGA 2328
D 2496 GGATTGGAATGCAAGAGTCTCAAACTGATTAATATTGGAAGAGA 2541

RESULT 5
AAS46128
ID AAS46128 standard; cDNA; 2609 BP.
XX
XX
AAS46128;
AC
XX
DT 18-DEC-2001 (first entry)
XX
XX
Human DNA encoding PRO polypeptide sequence #204.
DE
XX
XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW dog; cat; pig; goat; rabbit; tumour; neurosis; factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200168848-A2.
PD
XX
XX 20-SEP-2001.
PE
XX 28-FEB-2001; 2001MO-US06520.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05841.
PR 03-MAR-2000; 2000US-187202P.


```

Db 1520 GTACCATCTGTAGACCTCGGGCTGGAGCTCCGAGCTCCCTCATTTGGACCAAGGA 1579
OY 1563 CAGGAAAGAGAGAAATAACAGAGATAGAGAGCGGACAGAGCTTAGGCTATTCCTG 1622
Db 1580 CAGGAGAGAGAGAAATAACAGAGATAGAGAGCGGACAGAGCTTAGGCTATTCCTG 1639
OY 1623 CTGAAGCTTTCCCGAAGAGTCAAGCCCGAGTCTCTGACTCTGACTCTGACTCTGACTCTG 1682
Db 1640 CTGAAGCTTTCCCGAAGAGTCAAGCCCGAGTCTCTGACTCTGACTCTGACTCTGACTCTG 1699
OY 1683 CCTGTACCCCTGAGAGACTTCCACCCCGAGTCTCTGACTCTGACTCTGACTCTGACTCTG 1742
Db 1700 CCTGTACCCCTGAGAGACTTCCACCCCGAGTCTCTGACTCTGACTCTGACTCTGACTCTG 1759
OY 1743 ATCATGTCTCCCTGATAGAAATGTGTGATTTGATTTCAAGGCTTCAAGTATTTTCA 1802
Db 1760 ATCATGTCTCCCTGATAGAAATGTGTGATTTGATTTCAAGGCTTCAAGTATTTTCA 1819
OY 1803 CTATTTTCAAGAAATAGATAGATTTGCGGGGCTGAGTCTATTTCAAGAGACTGTG 1862
Db 1820 CTATTTTCAAGAAATAGATAGATTTGCGGGGCTGAGTCTATTTCAAGAGACTGTG 1879
OY 1863 AACAGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1922
Db 1880 AACAGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1939
OY 1923 CAAGAAGCCGGAGCTGGCGGGAACCTGGAGAGTCTGATTTCTTTTCCGTACACA 1982
Db 1940 CAAGAAGCCGGAGCTGGCGGGAACCTGGAGAGTCTGATTTCTTTTCCGTACACA 1999
OY 1983 GAGAGGCTATGTAAACAAACAGAGATCGAAGGGTTTGTAGAGATGTGTTTCA 2042
Db 2000 GAGAGGCTATGTAAACAAACAGAGATCGAAGGGTTTGTAGAGATGTGTTTCA 2059
OY 2043 AACCATGCTGTGTATTTCAACCAATAAAGAGTTGCTGCTCTTAATTTGATTAAC 2102
Db 2060 AACCATGCTGTGTATTTCAACCAATAAAGAGTTGCTGCTCTTAATTTGATTAAC 2119
OY 2103 GGTATATCTGTCTGTATTTGATTTGATTTTAAATAATGTGATAGATTTCTCTG 2162
Db 2120 GGTATATCTGTCTGTATTTGATTTGATTTTAAATAATGTGATAGATTTCTCTG 2179
OY 2163 GAAAGGCTTCAAGACATGCTATGTCTGTCTTCCAAACCCAGTCTCTCTCTCTCTCT 2222
Db 2180 GAAAGGCTTCAAGACATGCTATGTCTGTCTTCCAAACCCAGTCTCTCTCTCTCTCT 2239
OY 2223 AGCCAGGTTTCTTGTAGAGACCCCTTAATCTGTCTCTCTTAAATTTTAAACCAAT 2282
Db 2240 AGCCAGGTTTCTTGTAGAGACCCCTTAATCTGTCTCTCTTAAATTTTAAACCAAT 2299
OY 2283 GGATTTGATGTCAGAGCTCTCAAACTGATTAATTTTGAAGAGA 2328
Db 2300 GGATTTGATGTCAGAGCTCTCAAACTGATTAATTTTGAAGAGA 2345

```

RESULT 4
AA05359
ID AAX05359 standard; DNA: 2550 BP.
XX

XX AAX05359:
XX 07-MAY-1999 (first entry)
XX Human extracellular matrix protein (ECMP)-1 encoding DNA.
XX DE Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
XX KW Immune disorder; human; ss.
XX OS Homo sapiens.
XX FH Key
FT CDS Location/Qualifiers
FT 382..1728
FT /*tag= a

```

FT XX /product= "ECMP-1"
XX PN W09900410-A2.
XX PD 07-JAN-1999.
XX PF 23-JUN-1998; 98WO-US13012.
XX PR 27-JUN-1997; 97US-0884072.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Corley NC, Guegler KJ;
XX DR WPI: 1998-095674/08.
XX DR P-PSDB: AAW94281.
XX PT New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX PT useful in the diagnosis, prevention and treatment of immune
XX PT disorders and cancer
XX PS Claim 3; Fig 1A-G; 79pp: English.
XX CC This DNA encodes a human extracellular matrix protein (ECMP)-1. Host
XX CC cells containing a vector comprising the ECMP-1 nucleic acid are used
XX CC for the recombinant production of the protein. ECMP-1 and its
XX CC (antibodies, are useful in the diagnosis, prevention, and treatment
XX CC of cancer and immune disorders.
XX SQ Sequence 2550 BP; 623 A; 690 C; 596 G; 641 T; 0 other:
XX
XX Query Match 99.8%; Score 2324.4; DB 20; Length 2550;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 3 CCGGGGCTCTCCCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62
Db 216 CCGGGGCTCTCTCCCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 275
OY 63 GCGAGCCCGGAGGCGGAGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 122
Db 276 GCGAGCCCGGAGGCGGAGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 335
OY 123 CGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
Db 336 CGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 395
OY 183 AAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
Db 396 AAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 455
OY 243 AACGTGACGAGATGCTTGAACCTGATGCGGAGTCCAGAGAGAGTGTGATTTGATTA 302
Db 456 AACGTGACGAGATGCTTGAACCTGATGCGGAGTCCAGAGAGAGTGTGATTTGATTA 515
OY 303 ATGCCGAACCATCCCGGAGGCTTCCGAGAGACATGATGTGTTAAACCAAAATGCGCG 362
Db 516 ATGCCGAACCATCCCGGAGGCTTCCGAGAGACATGATGTGTTAAACCAAAATGCGCG 575
OY 363 GTATTTATGATTTCCCGGAGCAAACTGTGTATGAGGGGCTTACTCGAACCCCTACTC 422
Db 576 GTATTTATGATTTCCCGGAGCAAACTGTGTATGAGGGGCTTACTCGAACCCCTACTC 635
OY 423 GAGCCCTTACGACGATGCTGACCCGAGAGCTGCGGAGAGATGATGTGTTAAACCAAAATGCGCG 482
Db 636 GAGCCCTTACGACGATGCTGACCCGAGAGCTGCGGAGAGATGATGTGTTAAACCAAAATGCGCG 655
OY 483 CAGCATCTCCAGGCTCTTATATGCGGCTTTGATACGAGATGATGATGATGATGATGATGATGATGATGAT 542
Db 696 CAGCATCTCCAGGCTCTTATATGCGGCTTTGATACGAGATGATGATGATGATGATGATGATGATGATGAT 755
OY 543 TGTGATGTGAGAGATGTGACACAGATTTCCACAGTGCACACCCACGATCTGAC 602

```


QY	723	TGATTCCTATCTCTGTACATCCAAACCCGGTTTAACTCAATGAGATGGAAGTCTGTG	782
Db	740	TGATTCCTATCTCTGTACATCCAAACCCGGTTTAACTCAATGAGATGGAAGTCTGTG	799
QY	783	CCAAGATGTGAACGAGTGTGACCACCGAAGACCCCTGGGTGCAAACTGCTGTCAACACTTA	842
Db	800	CCAAGATGTGAACGAGTGTGACCACCGAAGACCCCTGGGTGCAAACTGCTGTCAACACTTA	859
QY	843	CGGCTCTTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGGAAGATGGCCTTCATTG	902
Db	860	CGGCTCTTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGGAAGATGGCCTTCATTG	919
QY	903	CAGTGAATGAGACAGTGTGACAGCTCTGTGAGATTCCTCTGCGCAATGAGTGTGGAACCA	962
Db	920	CAGTGAATGAGACAGTGTGACAGCTCTGTGAGATTCCTCTGCGCAATGAGTGTGGAACCA	979
QY	963	GCCCCGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGCTGATGCAACCGAAG	1022
Db	980	GCCCCGACATATCTTCTGCTCTGCCCTCCAGGCTACATCTCTGCTGATGCAACCGAAG	1039
QY	1023	CTGGCAAGACATCAAGAAATGTGAGCAAGGAACACACGCGCAACCTGCGACGAGAGGTG	1082
Db	1040	CTGGCAAGACATCAAGAAATGTGAGCAAGGAACACACGCGCAACCTGCGACGAGAGGTG	1099
QY	1083	CTACAAATTTACAAAGGGGGCTTCAAAATGCATGACCCATCCGCTGTGAGAGACCTTATCT	1142
Db	1100	CTACAAATTTACAAAGGGGGCTTCAAAATGCATGACCCATCCGCTGTGAGAGACCTTATCT	1159
QY	1143	GAGATATAGATTAACGGCTGTATGTGTCGCGTGAACACCTGGCTGCGACAGACCGACC	1202
Db	1160	GAGATATAGATTAACGGCTGTATGTGTCGCGTGAACACCTGGCTGCGACAGACCGACC	1219
QY	1203	CTTTACCATCTTGTACCGGGACATGAGCGTGTGTCAGAGCGTCCGTTCCCGCTGACAT	1262
Db	1220	CTTTACCATCTTGTACCGGGACATGAGCGTGTGTCAGAGCGTCCGTTCCCGCTGACAT	1279
QY	1263	CTTTCCAAATGCAAGCCACGACCCGCTACCTGGGGGCTATTACATTTTCCAGATCAATTC	1322
Db	1280	CTTTCCAAATGCAAGCCACGACCCGCTACCTGGGGGCTATTACATTTTCCAGATCAATTC	1339
QY	1323	TGGGAATGAGGGGCGAGAAATTTTACATGCGGCAAGGGGCCCCATCAAGTGCACCCGTGT	1382
Db	1340	TGGGAATGAGGGGCGAGAAATTTTACATGCGGCAAGGGGCCCCATCAAGTGCACCCGTGT	1399
QY	1383	GATGACACAGCCCCATCAAGAGGGGCCGGGAAATTCACAGCTGACCTTGGAAATGATCACTGT	1442
Db	1400	GATGACACAGCCCCATCAAGAGGGGCCGGGAAATTCACAGCTGACCTTGGAAATGATCACTGT	1459
QY	1443	CAACACATGTCATCAACTTCAGAGGCACTCCGTATCCGACTGGGGATATATGTGTGCGCA	1502
Db	1460	CAACACATGTCATCAACTTCAGAGGCACTCCGTATCCGACTGGGGATATATGTGTGCGCA	1519
QY	1503	GTAACCACTTCTGAGACCTCGGGCTGTGAGACCTTCGAGCGTGCCTCTCATATGCGACCAAGGA	1562
Db	1520	GTAACCACTTCTGAGACCTCGGGCTGTGAGACCTTCGAGCGTGCCTCTCATATGCGACCAAGGA	1579
QY	1563	CAGAGGAAGAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTGTG	1622
Db	1580	CAGAGGAAGAGAGAAATTAACAGAGAAATGAGAGCGACACAGACGTTAGGCATTTCTGTG	1639
QY	1623	CTGAACCTTTTCCCGGAGAGAGTCAACCCCGACTTCTCTGACCTCTCACTGATATTGGAGA	1682
Db	1640	CTGAACCTTTTCCCGGAGAGAGTCAACCCCGACTTCTCTGACCTCTCACTGATATTGGAGA	1699
QY	1683	CTGTGTACCTGTGACAGACTTGCCACCCCGAGTTCTATGATACAGTTATCAAAAAGATTT	1742
Db	1700	CTGTGTACCTGTGACAGACTTGCCACCCCGAGTTCTATGATACAGTTATCAAAAAGATTT	1759
QY	1743	ATCATTTGCTCCCGATAGAAAGATTTGGGATTTTCAAGGGCTTCAGTTTATTTTCCA	1802
Db	1760	ATCATTTGCTCCCGATAGAAAGATTTGGGATTTTCAAGGGCTTCAGTTTATTTTCCA	1819

OY	1803	CTATTTTCAAAGAAAATATGATTAGATTGTCGGGGGGTCGATCTATGTTCTCAAAACGTG	1862
Db	1820	CTATTTTCAAAGAAAATGATTAGATTGTCGGGGGGTCGATCTATGTTCTCAAAACGTG	1879
OY	1863	AACAGCTTGCTGACATTCTTCACCTCTTCACCTCTTCACCTGTTACTGCTTG	1922
Db	1880	AACAGCTTGCTGACATTCTTCACCTCTTCACCTCTTCACCTGTTACTGCTTG	1939
OY	1923	CAAGAACCCGGAGAGCTGGCGGGGAAACCTGGGATAGCTATTTGCTTTTGGCTAACAA	1982
Db	1940	CAAGAACCCGGAGAGCTGGCGGGGAAACCTGGGATAGCTATTTGCTTTTGGCGTACAA	1999
OY	1983	GAGAAAGCTATGATTAACCAAAACACAGCAGATGCGAAGGTTTTAGAAATGTGTTCAA	2042
Db	2000	GAGAAAGCTATGATTAACCAAAACACAGCAGATGCGAAGGTTTTAGAAATGTGTTCAA	2059
OY	2043	AACCATGCGCTGATTTTTCACACATTAAGAAGATTCAGTGTCTCTTAATTTGTATTAAC	2102
Db	2060	AACCATGCGCTGATTTTTCACACATTAAGAAGATTCAGTGTCTCTTAATTTGTATTAAC	2119
OY	2103	GATTAAATCTGCTGTGTTCAATTTTGAGATTTTAAAAAATATGCTAGATAATTCCTTC	2162
Db	2120	GATTAAATCTGCTGTGTTCAATTTTGAGATTTTAAAAAATATGCTAGATAATTCCTTC	2179
OY	2163	GAAAGGCTTGAGACACATGCTATGTTGTGCTTCCCAAACCCAGTCCCTCTGCATTTT	2222
Db	2180	GAAAGGCTTGAGACACATGCTATGTTGTGCTTCCCAAACCCAGTCCCTCTGCATTTT	2239
OY	2223	AGCCAGTGTGTTCTTTGAGGACCCCTTAATCTGTGCTTTCTTTAGAAATTTTACCACAT	2282
Db	2240	AGCCAGTGTGTTCTTTGAGGACCCCTTAATCTGTGCTTTCTTTAGAAATTTTACCACAT	2299
OY	2283	GGATTGGAATGCAAGGTCGCCAAACGATTAATTAATTTTGAAGA GA 2328	
Db	2300	GGATTGGAATGCAAGGTCGCCAAACGATTAATTAATTTTGAAGA GA 2345	

XX	RESULT 3
XX	AAZ39892
ID	AAZ39892 standard; cDNA: 2362 BP.
XX	
AC	AAZ39892;
XX	
DI	21-FEB-2000 (first entry)
XX	
DE	cDNA sequence of the human secreted protein AK647.
XX	
KW	AK647; aortic tissue development; smooth muscle cell modulator; SCID;
KW	nutritional supplement; vasculogenesis; embryonic development; infection;
KW	cytokine activity; cell proliferation; cell differentiation; detect; HIV;
KW	immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
KW	wound healing; restenosis; atherosclerosis; drug screen; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	185..1532
FT	/*tag=a
FT	/product= AK647
FT	/note="Secreted protein"
XX	
PN	W03960125-A2.
XX	
PD	25-NOV-1999.
XX	
PF	18-MAY-1999; 99WO-US10931.
XX	
PR	19-MAY-1998; 98US-0081002.
PR	21-MAY-1998; 98US-0083002.
XX	
PA	(GENM) GENETICS INST INC.
XX	
PI	Jacobs K, McCoy JM, Racie L, Lavallic E, Treacy M, Evans C;

```

OY 1801 CACTATTTCAGAAAGAAATAGATTAGTTGGGGGGCTGAGCTATGTTCAAGACTG 1860
    |||||||
DB 1801 CACTATTTCAGAAAGAAATAGATTAGTTGGGGGGCTGAGCTATGTTCAAGACTG 1860
OY 1861 TGAACAGCTGTGCTGCTACCTCTTCCACTCTCTCTCAGCTGTTACTGCTT 1920
    |||||||
DB 1861 TGAACAGCTGTGCTGCTACCTCTTCCACTCTCTCTCAGCTGTTACTGCTT 1920
OY 1921 TGCAGAGAGCCGGGAGCTGGGGGAAACCTGGGAGTACGTAGTTGCTTTTCCGTACA 1980
    |||||||
DB 1921 TGCAGAGAGCCGGGAGCTGGGGGAAACCTGGGAGTACGTAGTTGCTTTTCCGTACA 1980
OY 1981 CAGAGAGGCTATGTATTAACAAACACAGAGATGCAAGGGTTTGTAGAGATGTGTTCC 2040
    |||||||
DB 1981 CAGAGAGGCTATGTATTAACAAACACAGAGATGCAAGGGTTTGTAGAGATGTGTTCC 2040
OY 2041 AAACCCATGCTGTGATTTTCAACATTAAGAGATTGCTGCTTAAATTTGTATA 2100
    |||||||
DB 2041 AAACCCATGCTGTGATTTTCAACATTAAGAGATTGCTGCTTAAATTTGTATA 2100
OY 2101 ACGGTTTAACTGCTGCTTCTGCTATTTTGAATTTTAAATAATATGCTGAGATTGCT 2160
    |||||||
DB 2101 ACGGTTTAACTGCTGCTTCTGCTATTTTGAATTTTAAATAATATGCTGAGATTGCT 2160
OY 2161 TCGAAGAGCCTTCAGACATGCTATGTTCTGCTCCCAAGCCAGTCTCTCCATTT 2220
    |||||||
DB 2161 TCGAAGAGCCTTCAGACATGCTATGTTCTGCTCCCAAGCCAGTCTCTCCATTT 2220
OY 2221 TTAGCCCAAGTGTCTTCTTGAAGACCCCTTAATCTTCTTCTTGAATTTTACCCAA 2280
    |||||||
DB 2221 TTAGCCCAAGTGTCTTCTTGAAGACCCCTTAATCTTCTTCTTGAATTTTACCCAA 2280
OY 2281 TTGATGTGAATGACAGAGGCTTCCAAACTGATTAATTTTGAAGAGA 2328
    |||||||
DB 2281 TTGATGTGAATGACAGAGGCTTCCAAACTGATTAATTTTGAAGAGA 2328

RESULT 2
AA07567
ID AA07567 standard; cDNA; 2362 BP.
AC AA07567;
XX
XX 21-JUN-1999 (first entry)
DT
XX
XX Homo sapiens fetal kidney clone AK647 secreted protein gene.
DE
XX Secreted protein; fetal kidney; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 186..1532
FT CDS /*tag= "a
FT /*product= "secreted protein"
PN
XX
XX MO9900405-A1.
XX
XX 07-JAN-1999.
XX
XX 29-JUN-1998; 98WO-US13530.
XX
XX 30-JUN-1997; 97US-0885610.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX Agostino MJ, Evans C, Jacobs K, Lavallie ER, Mccoy JM;
PI Merberg D, Racie LA, Treacy M;
XX
XX WPI; 1999-095671/08.
DR
XX P-PSDB; AAW95709.
XX

```

```

PT New polynucleotides encoding secreted human proteins - are derived
PT from fetal kidney or adult retina cDNA libraries, used as, e.g.
PT potential vaccines
PS
XX Disclosure: Pages 51-52; 76pp; English.
XX
XX The sequence is that encoding a secreted protein from a human fetal
XX kidney clone AK647. Such a sequence is predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemoclastic/chemokine activity, hemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, and tumour
XX inhibition activity. It is also stated to be useful for gene
XX therapy.
SQ
Sequence 2362 BP; 580 A; 650 C; 548 G; 584 T; 0 other:
Query Match 99.9%; Score 2326; DB 20; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 CCCGGGCGTCTCCCGCTGCTCTCTCCAGCACTGCTGCGCCCTCTGGAATAAACACC 62
    |||||||
DB 20 CCCGGGCGTCTCCCGCTGCTCTCTCCAGCACTGCTGCGCCCTCTGGAATAAACACC 79
OY 63 GCGAGCCCGGAGGGCCCAAGAGAGGCGGAGCGTCCGAGTCTCCGGGGGTCCCGCCG 122
    |||||||
DB 80 GCGAGCCCGGAGGGCCCAAGAGAGGCGGAGCGTCCGAGTCTCTCCGGGGGTCCCGCCG 139
OY 123 CGAGCTTCTCTCGCTTCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
    |||||||
DB 140 CGAGCTTCTCTCTCGCTTCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 199
OY 183 AAGGATACCTACTGTTACCATTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
    |||||||
DB 200 AAGGATACCTACTGTTACCATTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259
OY 243 ACAGTGCAGATGAGCTTGTGACCTGATGCGCAGTACAGAGAGTGTAGATATTGTATGA 302
    |||||||
DB 260 ACAGTGCAGATGAGCTTGTGACCTGATGCGCAGTACAGAGAGTGTAGATATTGTATGA 319
OY 303 ATGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 362
    |||||||
DB 320 ATGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTAACCAAAATGGCGG 379
OY 363 GTATTATGACATTTCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC 422
    |||||||
DB 380 GTATTATGACATTTCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTC 439
OY 423 GACCCCTACTCAGGTCGTACCCAGACAGTGGCCCAACCACTCTCAGCTCTCAACTATTC 482
    |||||||
DB 440 GACCCCTACTCAGGTCGTACCCAGACAGTGGCCCAACCACTCTCAGCTCTCAACTATTC 499
OY 483 CACGATCTCAGAGCCTCTTATATGCGCTTTGGATACAGATGATGAAGAAACCAACATG 542
    |||||||
DB 500 CACGATCTCAGAGCCTCTTATATGCGCGTTTGGATACAGATGATGAAGAAACCAACATG 559
OY 543 TGTGATGTGAGCAGTGTGCAACAGATTTCCACCACTGTGCAACCCCAACCAATCTGCAT 602
    |||||||
DB 560 TGTGATGTGAGCAGTGTGCAACAGATTTCCACCACTGTGCAACCCCAACCAATCTGCAT 619
OY 603 CAATACGAAGCGGGTACACCTGCTCCGCAACCGAGGATATTTGGCTTTGGAAGGCCA 662
    |||||||
DB 620 CAATACGAAGCGGGTACACCTGCTCCGCAACCGAGGATATTTGGCTTTGGAAGGCCA 679
OY 663 GTGCTTAGACATTTGATGATGTGCTATGTTACTGTCAGACAGCTCTGTGCGAATGTTC 722
    |||||||
DB 680 GTGCTTAGACATTTGATGATGTGCTATGTTACTGTCAGACAGCTCTGTGCGAATGTTC 739

```

DR P-PSDB; AAY54989.

XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
PT myoma -
PS
XX Claim 4, 5; Page 76-80; 87pp; Japanese.

XX This sequence encodes the human A55 protein of the invention. The protein
CC can be used for the treatment of diseases due to abnormal proliferation
CC of smooth muscle. The polypeptides can be used according their inhibition
CC of the proliferation of vascular smooth muscle cells, particularly in
CC treating arteriosclerosis or re-narrowing by vascular endothelial
CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
CC or myoma, haemopoietic cell-regulatory activity, cytokine activity,
CC tissue generation/repairation activity, actin/inhibin activity, taxis
CC and chemotaxis activity, blood coagulation/thrombotic activity,
CC receptor/ligand activity, cachectin/tumour metastasis inhibiting activity;
CC Note: The open reading frame of this sequence, as well as the mature
CC protein coding sequence are specifically claimed in claim 4.

XX Sequence 2328 BP; 560 A; 646 C; 541 G; 581 T; 0 other:

Query Match 100.0%; Score 2328; DB 21; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCCGGCGCTCCCGCTCTCTCCACAGATCGTCGCGCCCTCTGGAATAAACAAC 60
DB 1 GACCCGGCGCTCCCGCTCTCTCCACAGATCGTCGCGCCCTCTGGAATAAACAAC 60
QY 61 CCGCGAGCCCGGAGGCGCCAGAGAGGCGAGCGTCCCGAGCTCTCGGGGGTCCCGCC 120
DB 61 CCGCGAGCCCGGAGGCGCCAGAGAGGCGAGCGTCCCGAGCTCTCGGGGGTCCCGCC 120
QY 121 CCGGAGCTTCTCTCGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 121 CCGGAGCTTCTCTCGCTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 181 AAAAGATATCTACTGTATACATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
DB 181 AAAAGATATCTACTGTATACATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 GCACAGTGCAGATGGCTTGAACCTGGATGCGAGGAGATGTTTATATATATAT 300
DB 241 GCACAGTGCAGATGGCTTGAACCTGGATGCGAGGAGATGTTTATATATATATATAT 300
QY 301 GAATGCCGAACATCCCGGAGGCGTCCGAGAGAGATGATGTGTATACCAAAATGCG 360
DB 301 GAATGCCGAACATCCCGGAGGCGTCCGAGAGAGATGATGTGTATACCAAAATGCG 360
QY 361 GGGTATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGGCGCTTACTCGAACCCCTAC 420
DB 361 GGGTATTTATGATTCCTCCGAGCAAAACCTGTGTATGAGGCGCTTACTCGAACCCCTAC 420
QY 421 TCGACCCCTACTACAGGTCCGTACCAAGACGTGCCACACAGTCTGAGCTCAAACTAT 480
DB 421 TCGACCCCTACTACAGGTCCGTACCAAGACGTGCCACACAGTCTGAGCTCAAACTAT 480
QY 481 CCCAGATCTCAGGCGCTTATATATGCGCTTGTGATACAGATGATGAAAGCAACCA 540
DB 481 CCCAGATCTCAGGCGCTTATATATGCGCTTGTGATACAGATGATGAAAGCAACCA 540
QY 541 TGTGTGATGTGACAGATGTGCAACAGATTCACACAGTGCACACCCACAGATGTGC 600
DB 541 TGTGTGATGTGACAGATGTGCAACAGATTCACACAGTGCACACCCACAGATGTGC 600
QY 601 ATCAATATGAAGGCGGTACACTGCTCTGACACGAGATATGTGCTTGTGAAGGC 660
DB 601 ATCAATATGAAGGCGGTACACTGCTCTGACACGAGATATGTGCTTGTGAAGGC 660
QY 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720

DB 661 CAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 CTTGATCTTATTTTGTATGCAACCTGGTTTACCCTAATAGAGATGAGAGTCT 780
DB 721 CTTGATCTTATTTTGTATGCAACCTGGTTTACCCTAATAGAGATGAGAGTCT 780
QY 781 TGGCAATGATGAGAGAGTGTGCCAGAGAACCCCTGGTGTCAAAACCGGTCAACAC 840
DB 781 TGGCAATGATGAGAGAGTGTGCCAGAGAACCCCTGGTGTCAAAACCGGTCAACAC 840
QY 841 TAGGCTCTTATCTGCGCTGTGACCCAGATATGAACTTGAAGAGATGAGAGTCT 900
DB 841 TAGGCTCTTATCTGCGCTGTGACCCAGATATGAACTTGAAGAGATGAGAGTCT 900
QY 901 TGCAGTATATGACAGAGTGTGACCTTCTGTGATTCCTTGTCCAAATATAGTGTGAAC 960
DB 901 TGCAGTATATGACAGAGTGTGACCTTCTGTGATTCCTTGTCCAAATATAGTGTGAAC 960
QY 961 CAGCCGGGCAATCTCTGCTCTGCGCTCCAGCTACATCTGCTGTGATGCAACCGA 1020
DB 961 CAGCCGGGCAATCTCTGCTCTGCGCTCCAGCTACATCTGCTGTGATGCAACCGA 1020
QY 1021 AGCTGCCAAGATCAACGAATGTGAGCAGAGAACCAACAGTGCACCTGACAGACG 1080
DB 1021 AGCTGCCAAGATCAACGAATGTGAGCAGAGAACCAACAGTGCACCTGACAGACG 1080
QY 1081 TGTCAATATTTACAAAGGGGCTTCAATATGATCGAACCCATCCGCTGTAGAGACCTTAT 1140
DB 1081 TGTCAATATTTACAAAGGGGCTTCAATATGATCGAACCCATCCGCTGTAGAGACCTTAT 1140
QY 1141 CTGAGATAGATGATTAACCGCTGTATGTCTCTGTGAAACCTGTGTCAGAGACCA 1200
DB 1141 CTGAGATAGATGATTAACCGCTGTATGTCTCTGTGAAACCTGTGTCAGAGACCA 1200
QY 1201 CCTTTACATCTTGTACCGGAGATGAGCTGTGTGAGAGCGTCTCTCCCGCTGAC 1260
DB 1201 CCTTTACATCTTGTACCGGAGATGAGCTGTGTGAGAGCGTCTCTCCCGCTGAC 1260
QY 1261 ATCTTCCAAATGCAAGCCACACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAAA 1320
DB 1261 ATCTTCCAAATGCAAGCCACACCCGCTACCTGGGGCTTATTAATTTTCCAGATCAAA 1320
QY 1321 TCTGGGAATGAGGCGCAGAAATTTTACATGGGCGCAACGGGGCCCATCAGTCCACCTG 1380
DB 1321 TCTGGGAATGAGGCGCAGAAATTTTACATGGGCGCAACGGGGCCCATCAGTCCACCTG 1380
QY 1381 GTGATGACAGGCCCCATCAAAAGGGCCCCGGAAATCCAGCTGAGCTTGAATAATGATCAGT 1440
DB 1381 GTGATGACAGGCCCCATCAAAAGGGCCCCGGAAATCCAGCTGAGCTTGAATAATGATCAGT 1440
QY 1441 GTCAACACTGTATCAACTTGAAGAGGAGCTCCGTGTATCCGACTGCGGATATATGTGTG 1500
DB 1441 GTCAACACTGTATCAACTTGAAGAGGAGCTCCGTGTATCCGACTGCGGATATATGTGTG 1500
QY 1501 CAGTACCAATTTCTAGGCTCGGGCTGAGAGCTCCGACAGCTGCTCATTTGGACCAAG 1560
DB 1501 CAGTACCAATTTCTAGGCTCGGGCTGAGAGCTCCGACAGCTGCTCATTTGGACCAAG 1560
QY 1561 GACAGGAGAGAGAGAAATTAACAGAGAAATGAGAGGAGACAGACGTTAGGCAATTTCC 1620
DB 1561 GACAGGAGAGAGAGAAATTAACAGAGAAATGAGAGGAGACAGACGTTAGGCAATTTCC 1620
QY 1621 TGTGGAAGCTTTCCCGGAGAGTGAAGCCCGGACTTCTGACCTGCTCATTTATTTGA 1680
DB 1621 TGTGGAAGCTTTCCCGGAGAGTGAAGCCCGGACTTCTGACCTGCTCATTTATTTGA 1680
QY 1681 GACCTGTACACCTGACAGACTTGCACCCAGTCTATATATATATATATATATATATAT 1740
DB 1681 GACCTGTACACCTGACAGACTTGCACCCAGTCTATATATATATATATATATATATATAT 1740
QY 1741 TTAATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCACTTATTTTC 1800
DB 1741 TTAATCATTTGCTCCCTGATAGAGATTTGTTGTAATTTTCAAGGCTTCACTTATTTTC 1800

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 518.747 Seconds
(without alignments)
10106.379 Million cell updates/sec

Title: US-09-674-379a-12

Perfect score: 2328
Sequence: 1 gaccgcgcgcctcccccgtg.....tgattaatatttgagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2328	100.0	2328	21	AAZ40027
2	2326	99.9	2326	20	AAZ40027
3	2326	99.9	2326	21	AAZ39892
4	2324.4	99.8	2550	20	AAZ05359
5	2323.8	99.8	2609	22	AAZ46128
6	2322.8	99.8	2609	22	AAZ46128
7	2112.8	90.8	2509	20	AAZ37670
8	2085.4	89.6	2126	22	AAZ94505
9	1674.2	71.9	1720	19	AAV62432

10	1674.2	71.9	1720	24	ABK13627	Human cDNA encoding
11	1534.8	65.9	1538	22	AAK94129	Human full-length
12	1480.8	63.6	2233	21	AAZ39384	Smooth muscle prol
13	1480.8	63.6	2233	21	AAZ40029	Full length mouse
14	1449.8	62.3	2429	21	AAZ39387	Smooth muscle prol
15	1449.8	62.3	2429	21	AAZ40031	Full length mouse
16	1347	57.9	1347	24	ABU52528	Human up50 cDNA se
17	1222.2	52.5	1228	18	AAZ39380	Human extracellular
18	1113.6	47.8	1344	21	AAZ39383	Smooth muscle prol
19	1098.6	47.2	1383	21	AAZ39386	Smooth muscle prol
20	1056.2	45.4	1269	21	AAZ39385	Smooth muscle prol
21	1056.2	45.4	1269	21	AAZ39388	Smooth muscle prol
22	898	38.6	1254	21	AAZ61672	CDNA encoding rat
23	898	38.6	1254	22	AAZ92195	Skin cell cDNA, SE
24	896	38.5	1250	24	ABZ34757	Rat cDNA isolated
25	813.6	34.9	853	22	AAZ91556	Human cDNA 5'-end
26	813.6	34.9	853	22	AAZ93169	Human cDNA clone r
27	776.4	33.4	778	20	AAZ37673	Human PRO protein
28	636.4	27.3	680	21	AAZ44245	Human secreted exp
29	556	23.9	670	22	AAZ92195	Human cDNA 5'-end
30	556	23.9	670	22	AAZ93463	Human cDNA clone r
31	493	21.2	587	22	AAZ92416	Human cDNA 3'-end
32	448.4	19.3	519	18	AAZ88061	Partial cDNA clone
33	448.4	19.3	519	18	AAZ02141	Human secreted pro
34	445	19.1	541	22	AAZ92946	Human cDNA 3'-end
35	386	16.6	1480	21	AAZ14608	CDNA encoding a hu
36	386	16.6	1525	18	AAZ88974	Human extracellular
37	386	16.6	1531	20	AAZ60351	Human extracellular
38	386	16.6	1707	22	AAZ13771	Human cDNA sequenc
39	386	16.6	1875	24	AAZ51402	Human cDNA sequenc
40	386	16.6	1875	24	ABK40256	CDNA encoding huma
41	386	16.6	1994	21	AAZ18239	Lung cancer associ
42	386	16.6	2018	21	AAZ39800	Human S1-5 RCM-P-11
43	382	16.4	1874	21	AAZ58583	Human PR226 prote
44	374.4	16.1	377	20	AAZ37686	Human EST sequence
45	352.4	15.1	1358	21	AAA14616	CDNA encoding the

ALIGNMENTS

RESULT 1	AAZ40027	AAZ40027 standard; DNA; 2328 BP.
ID	AAZ40027	
XX	AAZ40027	
AC	AAZ40027	
XX		
DT	15-FEB-2000 (first entry)	
XX		
DE	Full length human A55 protein coding sequence.	
XX		
KW	A55 protein; human; smooth muscle proliferation; tissue generation;	
KW	vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;	
KW	percutaneous transluminal coronary angioplasty; blood coagulation; PBCA;	
KW	actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;	
XX	tumour metastasis inhibitor; ss.	
OS	Homo sapiens.	
XX		
PN	WO955864-A1.	
XX		
PD	04-NOV-1999.	
XX		
PF	28-APR-1999; 99WO-JP02284.	
XX		
PR	28-APR-1998; 98JP-0119731.	
XX		
PA	(ONOR) ONO PHARM CO LTD.	
XX		
PI	Honjo T, Tashiro K, Nakamura T;	
XX		
WPI	2000-038647/03.	


```

Db 90285 AAGTTTCAGTGTGCTTAATTTGTATTAACGGTTTAAATCTGCTGTTTCATTGAGTA 90344
OY 2133 TTTTAAAAAATATGTGTAGAAATTCCTTGAAAGGCTTCAGACATGCTATGTTCTG 2192
    |||||||
Db 90345 TTTTAAAAAATATGTGTAGAAATTCCTTGAAAGGCTTCAGACATGCTATGTTCTG 90404
    |||||||
OY 2193 TCTTCCCAAAACCCAGTCTCTCTCCATTTTAGCCCAAGTGTCTTCTTGAGAGACCCCTTAA 2252
    |||||||
Db 90405 TCTTCCCAAAACCCAGTCTCTCTCCATTTTAGCCCAAGTGTCTTCTTGAGAGACCCCTTAA 90464
    |||||||
OY 2253 TCTTCTTCTTCTTGTAGAAATTTTACCAATTTGATGTGAATGCAAGAGTCTCCAAACTGAT 2312
    |||||||
Db 90465 TCTTCTTCTTCTTGTAGAAATTTTACCAATTTGATGTGAATGCAAGAGTCTCCAAACTGAT 90524
    |||||||
OY 2313 TAAATATTGAGAGA 2328
    |||||||
Db 90525 TAAATATTGAGAGA 90540
    |||||||

```

Search completed: July 3, 2003, 15:34:27
 Job time : 6074.42 secs


```

QY 1953 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 2012
    |||||||
Db 16137 GGAGTAGCTAGTGTGCTTTTGGCTACACAGAGAGCGCTATGTAACCAACACAGCAGG 16078
QY 2013 ATCGAAGGCTTTTGAAGAAATGTGTTCAAAAACCATGCTGTGATTTTCAACCATAAAG 2072
    |||||||
Db 16077 ATCGAAGGCTTTTGAAGAAATGTGTTCAAAAACCATGCTGTGATTTTCAACCATAAAG 16018
QY 2073 AAGTTTCAGTGTGCTTTTGAATTTTGTATAGCGTTAATTCGTCTGTTCATTTTGAGCA 2132
    |||||||
Db 16017 AAGTTTCAGTGTGCTTTTGAATTTTGTATAGCGTTAATTCGTCTGTTCATTTTGAGCA 15958
QY 2133 TTTTAAAAAATATGCTGAGAAATTCCTGGAAGGCTTCAGACATGCTATGTCTG 2192
    |||||||
Db 15957 TTTTAAAAAATATGCTGAGAAATTCCTGGAAGGCTTCAGACATGCTATGTCTG 15898
QY 2193 TCTTCCCAACCCAGTCTCTCTGCAATTTTGAAGGCTTTTCTTTGAGAGACCCCTTAA 2252
    |||||||
Db 15897 TCTTCCCAACCCAGTCTCTCTGCAATTTTGAAGGCTTTTCTTTGAGAGACCCCTTAA 15838
QY 2253 TCTTGTCTTCTTGAATTTTACCCCAATGGATTGGAATGACAGAGCTCCCAACTGAT 2312
    |||||||
Db 15837 TCTTGTCTTCTTGAATTTTACCCCAATGGATTGGAATGACAGAGCTCCCAACTGAT 15778
QY 2313 TAAATATTGAAGAGA 2328
    |||||||
Db 15777 TAAATATTGAAGAGA 15762

```

RESULT 15

```

LOCUS 175210 bp DNA linear PRI 25-JUN-2001
CNS07EG5
DEFINITION Human chromosome 14 DNA sequence BAC R-818K5 of library RPI-11
from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL590328.2 GI:14571668
VERSION AL590328
KEYWORDS HTG: HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 175210)
Hellwig,R., Pellet,U.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brotlier,P., Catillo,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissbach,J.
Sequencing of the human chromosome 14

```

```

TITLE Unpublished
JOURNAL 2 (bases 1 to 175210)
REFERENCE Genoscope.
AUTHORS Direct Submission
JOURNAL Submitted (25-JUN-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr

```

COMMENT

```

On Jun 26, 2001 this sequence version replaced gi:13509295.
Web : www.genoscope.cns.fr)
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Sequefgenoscope.cns.fr

```

```

The following BAC sequence is oriented from the r7 to the sp6 end.
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.46x in Q20 bases; sum-of-contigs

```

```

Overall quality chart :
Range : bases
0 - 9 : 52
10 - 19 : 268
20 - 29 : 576

```

```

30 - 39 : 1644
40 - 49 : 7554
50 - 59 : 9152
60 - 69 : 10935
70 - 79 : 21682
80 - 89 : 43918
90 - 99 : 79423
-----
Percentage of bases with a quality value >= 40 : 98 %.
FEATURES
Location/Qualifiers
source
1..175210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-818K5"
/clone_id="RPI-11"
BASE COUNT 46924 a 36746 c 38389 g 53145 t 6 others
ORIGIN

```

```

Query Match 41.9% Score 974.4; DB 9; Length 175210;
Best Local Similarity 99.9%; Pred. No. 3.3e-261;
Matches 975; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1353 GCAAGCGGCCCATCATGATGCGACCTGCTGATGACAGCGCCCATCAAAAGGCGCCGGGA 1412
    |||||||
Db 89565 GCAAGCGGCCCATCATGATGCGACCTGCTGATGACAGCGCCCATCAAAAGGCGCCGGGA 89624
QY 1413 AATCCAGCTGACTTGGAAATGATCATCTGTCAACACTGTCAATCAACTTCAGAGCGACTC 1472
    |||||||
Db 89625 AATCCAGCTGACTTGGAAATGATCATCTGTCAACACTGTCAATCAACTTCAGAGCGACTC 89684
QY 1473 CGTGATCGAGTGGGATATATGTGTGTCGACGATACCATCTGAGCGCTGGGCTGAGGCT 1532
    |||||||
Db 89685 CGTGATCGAGTGGGATATATGTGTGTCGACGATACCATCTGAGCGCTGGGCTGAGGCT 89744
QY 1533 CCGAGCGTGCCTCATATGCGACCAAGGACAGAGAGAGAGAAATTAACAGAGAAAT 1592
    |||||||
Db 89745 CCGAGCGTGCCTCATATGCGACCAAGGACAGAGAGAGAGAAATTAACAGAGAAAT 89804
QY 1593 GAGAGCGACACAGAGCTTGAAGCATTTCTGCTGAACGTTTCCCGAAGAGTACGCCGGA 1652
    |||||||
Db 89805 GAGAGCGACACAGAGCTTGAAGCATTTCTGCTGAACGTTTCCCGAAGAGTACGCCGGA 89864
QY 1653 CTTCCTGACCTGACCTGCTGATATTTGACAGCTGTACCTGTCAGAGACTTGGCACCCCA 1712
    |||||||
Db 89865 CTTCCTGACCTGACCTGCTGATATTTGACAGCTGTACCTGTCAGAGACTTGGCACCCCA 89924
QY 1713 GTTCTATGATACAGTTATCAAAAAGTATATCATTTGCTCCCTGATAGAGATTGTTGG 1772
    |||||||
Db 89925 GTTCTATGATACAGTTATCAAAAAGTATATCATTTGCTCCCTGATAGAGATTGTTGG 89984
QY 1773 TGAATTTCAAGGCTTCAGTTATTTCCATATTTTCAAAAGAAATGATAGTTTTC 1832
    |||||||
Db 89985 TGAATTTCAAGGCTTCAGTTATTTCCATATTTTCAAAAGAAATGATAGTTTTC 90044
QY 1833 GGGGGTCTGAGTCTATGTTCAAAAGCTGTGAACAGCTGTCTGTCACTTCAACCTCTTC 1892
    |||||||
Db 90045 GGGGGTCTGAGTCTATGTTCAAAAGCTGTGAACAGCTGTCTGTCACTTCAACCTCTTC 90104
QY 1893 CACTCTTCTCTACAGTGTGTTACGCTTGTGCAAAAGACCCGGAGCTGGCGGGAACCTTC 1952
    |||||||
Db 90105 CACTCTTCTCTACAGTGTGTTACGCTTGTGCAAAAGACCCGGAGCTGGCGGGAACCTTC 90164
QY 1953 GGAAGTACTGTTGCTTTTGGCTGTACACAGAGAGGCTATGTAACAAACACAGCAGAG 2012
    |||||||
Db 90165 GGAAGTACTGTTGCTTTTGGCTGTACACAGAGAGGCTATGTAACAAACACAGCAGAG 90224
QY 2013 ATCGAAGGCTTTTGAAGAAATGTGTTCAAAACCATGCTGTATTTTCAACCATAAAG 2072
    |||||||
Db 90225 ATCGAAGGCTTTTGAAGAAATGTGTTCAAAACCATGCTGTATTTTCAACCATAAAG 90284
QY 2073 AAGTTTCAGTGTGCTTTTGAATTTTGTATAGCGTTAATTCGTCTGTTCATTTTGAGTA 2132
    |||||||

```

QY 1672 ACTATTGACAGCTGTGACCTGACAGACTTGGCACCAGTTCCTATGATACAGTTAT 1731
 Db 1653 ACTATCGCGAGA-CTGTCACCTCTGTAAGAGACACCAATGCCCCAGTTCTATGATGCAAGTTAT 1711
 QY 1732 CAAAAAGATATATATGCTGCTCCCGATAGAGAGATGTTGGTGAATTTTCAAGCCCTTCA 1791
 Db 1712 CCAAAAGATATCTCCCTTG-ACCCTGATGAGTTGCAAGTATGATCTCCAAAGCCCTTCC 1770
 QY 1792 GTTATTTCACATATTTT-----CAAGAAAAATAGATTAGTTTCCGGGGGCTGAGT 1844
 Db 1771 ATTATTTCATATTTTATTTTAAAAAGAAAAATAGATTAGTTTCTGGGGTGAAGT 1830
 QY 1845 CTATGCTCAAGACGTGTAACGCTGCTGCTCACTGCTTCACTGCTTCACTGCTTCT 1904
 Db 1831 CTGTTTCAAGAAC-TCAGACAGCTTGTCTTCTGCTTCCCTTCTTCT-CTATATCT 1887
 QY 1905 CACTGTGTACTGCTTTGCAAA--GACCCGGAGAGCTGGCGGGAACCCCTGGAGTACATA 1962
 Db 1888 TGCTCATATGCTGCTTGGCAAAAGTCCATGAGCTCATGGAATGCTGGGATATAGTA 1947
 QY 1963 GTTTCCTTTTCCGTCACAGAGAGAGCTATGTAACAAACACAGAGAGTCAAGGCT 2022
 Db 1948 TTTTCTCTTCTGCTGCTGCTGCAAGAGCTATGGCAACAGCTCACAGCAGATC-AAAGCT 2006
 QY 2023 TTTTAGAGATGTGTTTCAAAACCATGCTGATTTTCAACATATAAGAGTTTACGT 2082
 Db 2007 TTTTAAAGATGCTATTTTAAATTCAGCTGCTGATTTTTCAGCATTAAGAGATTTAGTT 2066
 QY 2083 TGTCTTAAATTTGTAATTAAGGTTTAACTGCTGCTGCTTCAATTTTGAATTTTAAAA 2142
 Db 2067 GTTCTTAAATTTGTAATGAGTGTAACTTTCTGCTGCTTCAATTTTGAATTTT----- 2119
 QY 2143 ATATGTCGAGATTCCTCGAAAGGCTTCAGACATCATATGCTGCTTCTTCCCAA 2202
 Db 2120 -AAATGATAGATTCCTTCGAAAGGCTTCAGATCATATGCTGCTTCTTCAAC 2178
 QY 2203 CCAAGCTCTCTCTCAATTTTACCCAGTGTCTTCTTGGAGAGCCCTTAACTTCTGCTT 2262
 Db 2179 CTCACCTTCTCCCTGCTGCTTACCCCA--GTTTTATGAGAACCCCTTAACTATGCTTTC 2236
 QY 2283 TTTAGATTTTATACCAATTTGATTTGAAATGAGAGTCTCAAACTGATTTAAATTTG 2322
 Db 2237 TTAAGAGTTTATACCAACTGCTGTTGGAAGACAGAGTATCCGACTGATTAATTTG 2296
 QY 2323 AAGAGA 2328
 Db 2297 AAGAAA 2302
 RESULT 14
 AC007057/c 171901 bp DNA linear PRI 25-APR-1999
 DEFINITION Homo sapiens chromosome 14 clone containing genes for UP50 and
 AC007057
 AC007057
 VERSION AC007057.3 GI:4680440
 KEYWORDS HMG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 171901)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
 Shafer, T. and Hood, L.
 Sequencing of human chromosome 14 gene for UP50 and partial gene
 for Trip230, thyroid hormone receptor co-activator
 Unpublished
 2 (bases 1 to 171901)
 Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 Hicks, P., James, R., Loretz, C., Lasky, S., Madan, A., Ratcliffe, A.,
 Shafer, T. and Hood, L.

TITLE Direct Submission
 JOURNAL Submitted (11-MAR-1999) Multimegabase Sequencing Center, University
 of Washington, PO Box 357730, Seattle, WA 98195, USA
 REFERENCE 3 (bases 1 to 171901)
 AUTHORS Rowen, L., Madan, A., Qin, S., Abbasi, N., Dors, M., Dickhoff, R.,
 James, R., Loretz, C., Lasky, S., Madan, A., Prescott, S., Ratcliffe, A.,
 Shafer, T. and Hood, L.
 TITLE JOURNAL
 Submitted (25-APR-1999) Multimegabase Sequencing Center, University
 of Washington, PO Box 357730, Seattle, WA 98195, USA
 COMMENT On Apr 25, 1999 this sequence version replaced gi:4558634.
 FEATURES
 Location/Qualifiers
 1..171901
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /map="14q31"
 /clone="BAC 353N19"
 /clone_lib="RPC111"
 complement(16075..93737)
 /gene="UP50"
 /note="The cDNA for this gene is found in AF093118."
 complement(115917..171768)
 /gene="Trip230"
 /note="The cDNA for this gene is found in AF007321. The
 5' end of the gene is missing from this clone."
 BASE COUNT 51278 a 38116 c 35671 g 46836 t
 ORIGIN
 Query Match 41.9%; Score 974.4; DB 9; Length 171901;
 Best Local Similarity 99.9%; Pred. No. 3.3e-261;
 Matches 975; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1353 GCAAGGGGCCCCATATGTCACACCCCTGGTATGACACAGCCCATCAAGGGCCCCGGGA 1412
 Db 16737 GCAAAGGGGCCCCATATGTCACACCCCTGGTATGACACAGCCCATCAAGGGCCCCGGGA 16678
 QY 1413 AATCCAGTGGAGCTTGGAAATGATTCATGCAACAGTCAATCAACTTCAGAGCAGTTC 1472
 Db 16677 AATCCAGTGGAGCTTGGAAATGATTCATGCAACAGTCAATCAACTTCAGAGCAGTTC 16618
 QY 1473 CGTATCCAGTGGAGCTTGGATATGATGTCGACAGTACCCATTCGAGCCCTGGAGCCT 1532
 Db 16617 CGTATCCAGTGGAGCTTGGATATGATGTCGACAGTACCCATTCGAGCCCTGGAGCCT 16558
 QY 1533 CCGAGCCTCCCTCTCATTTGGCACCAAGGACAGAGAAAGAGAAATPACAGAGAGAT 1592
 Db 16557 CCGAGCCTCCCTCTCATTTGGCACCAAGGACAGAGAAAGAGAAATPACAGAGAGAT 16498
 QY 1593 GAGAGGACACAGAGCTTGGGATTTCCCTGTAAGCTTCCCGGAAGAGTACAGCCCGGA 1652
 Db 16497 GAGAGGACACAGAGCTTGGGATTTCCCTGTAAGCTTCCCGGAAGAGTACAGCCCGGA 16438
 QY 1653 CTTCCTGACTCTACCTGTAATTTGACAGCTGTCAACCTGACAGACTTCCACCCCA 1712
 Db 16437 CTTCCTGACTCTACCTGTAATTTGACAGCTGTCAACCTGACAGACTTCCACCCCA 16378
 QY 1713 GTTCTGATGATACAGTTATCAAAAAGATTTATCATTTGCTCCCTGATGAGAGATTGTTG 1772
 Db 16377 GTTCTGATGATACAGTTATCAAAAAGATTTATCATTTGCTCCCTGATGAGAGATTGTTG 16318
 QY 1773 TGAATTTTCAAGGCTTCAAGTTATTTCCACTATTTTCAAAAGAAATAGATTAGTTTGC 1832
 Db 16317 TGAATTTTCAAGGCTTCAAGTTATTTCCACTATTTTCAAAAGAAATAGATTAGTTTGC 16258
 QY 1833 GGGGCTCAGCTATGATTCAAAGAGCTGTAAAGCTTGTGCTACCTTCTACCTCTTC 1892
 Db 16257 GGGGCTCAGCTATGATTCAAAGAGCTGTAAAGCTTGTGCTACCTTCTACCTCTTC 16198
 QY 1893 CACTCCTTCTCAGCTGTGTTACTGTTGCAAAAGACCCGGAGAGCTGGCGGGAACCTTC 1952
 Db 16197 CACTCCTTCTCAGCTGTGTTACTGTTGCAAAAGACCCGGAGAGCTGGCGGGAACCTTC 16138

QY 1313 AGATCAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAACGGGCCCATCTAGT 1372
 DB 1464 AGATCAATCTGGCAACGAGGCGTCAAGAGTTCTATATGCGGCAACAGGGCTTACAGT 1523
 QY 1373 CCACCCCTGGTATGACACGCCCATCAAAAGGCCCGGGAATTCACCTGAGCTTGGAA 1432
 DB 1524 CCACCCCTGGTATGACACGCCCATCAAAAGGCCCGGGAATTCACCTGAGCTTGGAA 1583
 QY 1433 TGATCAGTGTCAACAGTGTCAACAGTGTCAACAGTGTCAACAGTGTCAACAGTGT 1492
 DB 1584 TGATCAGTGTCAACAGTGTCAACAGTGTCAACAGTGTCAACAGTGTCAACAGTGT 1643
 QY 1493 ATGTGTGCGAGTACCATCTCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1552
 DB 1644 ATGTGTGCGAGTACCATCTCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT 1703
 QY 1553 CACCAAGGACGAGGAG 1612
 DB 1704 CACCAAGGACGAGGAG 1761
 QY 1613 GCATTTCTGCTGAACGTTTCCCGGAAGAGTCAAGCCCGGAGCTTCTGAGCTCACT 1672
 DB 1762 ACCTTTCTGCTGAATATCTCTGAGGAGCATCAGCCCTAG-CATTTGACCCATATCTGTA 1820
 QY 1673 CTATTGACAGACCTGTCACCCCTGAGAGACTTCCACCCAGTTCTCTATGATACAGTATC 1732
 DB 1821 CTATTGACAGA-TGCTGACTGTGAGAGACACCCCTGAGCTCTCTATGATACAGTATC 1879
 QY 1733 AAAAAATATATTCATTTGCTCCCTGATAGAGATTTGTTGTTGAATTTTCAAGCCCTCAG 1792
 DB 1880 CAAAAATATATTCATTTGCTCCCTGATAGAGATTTGTTGTTGAATTTTCAAGCCCTCAG 1938
 QY 1793 TTTATTTTCCATATTTT---CAAGAAATATATTTAGTTTGGGGGGTGTGAGTCT-- 1846
 DB 1939 TTTATTTTCCATGTTTATTAAGAAATATGATTTGTTGTTGGGGTGTGAGTCTCTC 1998
 QY 1847 -ATGTCAAGACCTGTCACCCCTGAGAGACTTCCACCCAGTTCTCTATGATACAGTATC 1905
 DB 1999 GAAAGTTCAAAAGAGTGTGCTTGTCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2057
 QY 1906 ACTGTGTTACTGCTTGGCAA--GACCCGGGAGCTGGCGGGGAAACCTGGGAGTGTAG 1963
 DB 2058 GCTGATTTGCTCTTGGCAAAGTCTCTCATGGGCTCTGCGGAAATGCTGGGAAATAGCTAG 2117
 QY 1964 TTTGCTTTTGGGTACACAGAGAGAGGCTATTAACAAACACAGAGATCGAAGGCT 2023
 DB 2118 TTTGCTTTTGGGTATGCTGAGAGAGGCTATGAGAGACACACAGAGATCGAAGGCT 2176
 QY 2024 TTTAGAGATGTTTCAAAAGCCTGCTGATTTTCAACATCAATCAAAAGATTTCACT 2083
 DB 2177 TTTATAGAGTATTTTAAATCAATCTGTATTTTCAACATCAAAAGATTTTCACT 2236
 QY 2084 GTCTTAAATTTGATACGCTTAAATTTCTGTCTGTCTGTTCAATTTGAGTATTTTAA 2143
 DB 2237 TCTTTAAATTTGATAGAGTGTAACTTTTCAATTTTCAATTTTGAAGCTTTTAA-- 2293
 QY 2144 TATGCTGAGATTTCTGGAAGGCTTCAACATCAATGATTTGTTGAGAGCCCTTAACTTCT 2203
 DB 2294 ---GTGGTATGATTTCTTCAAAAGGCT-TCAGATACAGTGTATGTTAGTCTTCAAC 2349
 QY 2204 CCAGTCTCTCTCAATTTTGAAGGCTTCTTCTTGTGAGAGCCCTTAACTTCTTCTCT 2263
 DB 2350 TCATCTCTCTCTCAATTTTGAAGGCTTCTTCTTGTGAGAGCCCTTAACTTCTTCTCT 2406
 QY 2264 TTAGAATTTTCCCAATTTGATGAGATGAGAGATGAGATGAGATGAGATGAGATGAG 2323
 DB 2407 TTAGAATTTTCCCAATTTGATGAGATGAGAGATGAGATGAGATGAGATGAGATGAG 2466
 QY 2324 AGAGA 2328
 DB 2467 AGAAA 2471

RESULT 12
 AF112153
 LOCUS
 DEFINITION
 Rattus norvegicus developmental arteries and neural crest EGF-like protein mRNA, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 2499)
 Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Tanikawa, M., Furukawa, Y., Koduke, K., Tashiro, K., Lu, Z., Andon, N.L., Schaub, R., Matsumori, A., Sasayama, S., Chien, K.R. and Honjo, T.
 DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries
 J. Biol. Chem. 274 (32), 22476-22483 (1999)
 MEDLINE
 PUBMED
 9935779
 10428823
 REFERENCE
 2 (bases 1 to 2499)
 Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
 Direct Submission
 Submitted (09-DEC-1998) Medicine, University of California San Diego, 9500 Gilman drive, La Jolla, CA 92093-0613, USA
 JOURNAL
 TITLE
 FEATURES
 source
 Location/Qualifiers
 1. 2499
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 350. 1696
 /note="DANCE: contains 6 cbegf domains and 1 RGD motif; secreted protein; expressed in developmental aorta, neural crest cells, balloon injured vessels, atherosclerotic lesions"
 /codon_start=1
 /product="developmental arteries and neural crest EGF-like protein"
 /protein_id="AA041769.1"
 /db_xref="GI:5305675"
 /translation="MPELRILVTITLALMPLRPNMAOQCTNGFDLDRQTCGLDID ECHTIEBACDGMKVMQNGYLICPTNPVRYSPYSPSTYSPPAPAPVAPAS NYPTISPLVCRFGYOMDEGNOCVDVDEACDSHOQCTOICINEGVTCGTGYW LREGCLDIDECRYGCOOCANPAGSYSCCNCFNTINDGRSDVNECETEPVQTCVNTGSPICRDPETELBEDGIRHSDMECSPEFLCQHECVNPGSTYSCPPG YVLEENRSCQDINECHRNHTCTPDTCTVNLQSGFKCIDVCEPYLLIDNRMC PALNTGCRDPTILFRDMDVSGRSVPADIFQOATRYGAYVIFIKSGNEBREF YMRQTPISATLVMTTRPIKGPRIQDLDEMTVTVIVFRGSSVIRLITYSQYP"
 BASE COUNT
 584 a 669 c 595 g 649 t 2 others
 ORIGIN
 Query Match 63.48; Score 1476.6; DB 10; Length 2499;
 Best Local Similarity 82.7%; Pred. No. 0;
 Matches 1839; Conservative 0; Mismatches 356; Indels 30; Gaps 12;
 QY 112 GGTCCCGCGCGAGCTTTCTCTGCGCTTGCAGATCTCTCTGCGGCTTGGAGATG 171
 DB 294 GGTCCCGCGAGAGCCCTCTCTCAATATACCTCTCTCTCTCTCTCTCTCTCTCTCT 352
 QY 172 CCAGGATTAAGAGATCTACTGTTACATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 231
 DB 353 CCAGGATTAAGAGATCTACTGTTACATTTGCTGCTCTCTCTCTCTCTCTCTCTCTCT 412
 QY 232 AATGACAGGACAGTGTGACCAATGCTTTGACCTGATGCCCACTAGACAGTGTTA 291
 DB 413 AATGACAGGACAGTGTGACCAATGCTTTGACCTGATGCCCACTAGACAGTGTTA 472
 QY 292 GATATGATGAATGCGGAACCATCTCCGAGGCGCTGCGAGGAGACATATGTTTAC 351
 DB 473 GATATGATGAATGCGGACCATCTCTGAGGCTTGGCGGAGACATATGTTTAC 532

Db	2090	TCATCCCTTCTCGATCTTACCCAGTTTAC---GAGACCCCTTAATATGCTTCT	2146
OY	2264	TTAGATTTTACCCAAATTTGATTTGGAATGAGAGTCTCCAAATGATTAAATTTGA	2333
Db	2147	TAAGATTTTACCCAACTGCGTTGGAAGACAGAGTATCCAGACTGATTAAATTTTGA	2206
OY	2324	AGAG 2328	
Db	2207	AAAA 2211	
RESULT 11	AF112151	2478 bp	linear
LOCUS	AF112151		
DEFINITION	Mus musculus developmental arteries and neural crest EGF-like		
ACCESSION	AF112151		
VERSION	AF112151.1		
KEYWORDS	GI:5305670		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Nakamura,T., Ruiz-Lozano,P., Lindner,V., Yabe,D., Taniwaki,M., Furukawa,Y., Kobuke,K., Tashiro,K., Lu,Z., Andon,N.L., Schaub,R., Matsumori,A., Sasayama,S., Chien,K.R. and Honjo,T.		
TITLE	DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries		
JOURNAL	J. Biol. Chem. 274 (32), 22476-22483 (1999)		
PMID	99357779		
REFERENCE	10428823		
AUTHORS	2 (bases 1 to 2478)		
TITLE	Nakamura,T., Yabe,D., Tashiro,K. and Honjo,T.		
JOURNAL	Submitted (09-DEC-1998) Medicine, University of California San Diego, 9500 Gilman drive, La Jolla, CA 92093-0613, USA		
FEATURES	Location/Qualifiers		
SOURCE	1..2478		
CDS	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	320..1666		
	/note="DANCE: contains 6 CBEGF domains and 1 RGD motif; secreted protein; expressed in developmental aorta, neural crest cells, balloon injured vessels, atherosclerotic lesions"		
	/codon_start=1		
	/product="developmental arteries and neural crest EGF-like protein"		
	/protein_id="AA041767.1"		
	/db_xref="GI:5305671"		
	/translation="MPGLRLTYVTLALMLPHGNAQOQCTNGFDLDRSGOGLDIDECRPIPEACRDMGCVONGVGLDIPETNIVRGRPYNPSTSGPYAPAPVPVMSNPILPSRPLRCRFGYOMDEKQCDVDEKQCPDRGCPYQICNTBEGYACSCIDGWLLEKQCDIDIBECRIGTQOQALCANPVSYSCTCNPGFTLNDGRSCDVANECETENPCVQTCNVTVGSFTCRCDPGELEBDGHSMDDECSFSEFLQHECVNPGSYFSCSPGQYLLADNRSCODINECHRNHTCTSLQTCYNLQGGFCIDPISCEEPYLLIGENRCMPAERTSCRDQPFILYRDMVSGRSVPADIFQWATRYGAYIFQISGNGRGEFYMRGTGISATLVMTTRPIKGRPDIDOLEMTVNTVINFGSSVIRLRITVSGYPP"		
BASE COUNT	604 a 656 c 575 g 643 t		
ORIGIN			
Query Match	63.7%; Score 1481.8; DB 10; Length 2478;		
Best Local Similarity	82.8%; Pred. No. 0;		
Matches 1842; Conservative	0; Mismatches 357; Indels 26; Gaps 12;		
OY	113	GTCCCGCCGCGAGACTTCTTCTGCGCTTGCATCTCTCTCGCGGCTTTGGACATGC	172
Db	264	GTCCACACCGCGAGACCTGCGCTTCTCTCGCGGCTTCTCTCGCGGCTTTGGATATGC	323
OY	173	CAGCAATTAAGATATCTACTCTTACCATCTGGCTCTCTGCTTCCAAAGCCCTTGGA	232

Dp	324	CAGAGTTTAAAAAGGATACTCACTGTTTACACATCTTGGCACTGTGGGTTCCACATCTCGGGA	383
Oy	233	ATGCACAGGCACTGTCACAGATGGCTTTGACCTGGATGCCAGTCCAGACAGTCTTAA	292
Dp	384	ATGCACACACAGTGCACAAAGGGCTTTGACCTGGACCGGCACTGACAGACAGTCTCTAG	443
Oy	293	ATATTGATGATGCCGAAACCATCCCCGAGGGCTGGCCGAGGAGACATGATGTGTATAAC	352
Dp	444	ATATTGATGAATGGCCGGACCATCTCTGAGGGCTTGTCTGGGGACATGATGTGTCAACC	503
Oy	353	AAAATGGGGGATATTATGACATTTCCCGGACAAACCTGTGTATGAGAGGCGCTCTGGA	412
Dp	504	AGATGGGGGGATTTTGTGCATCTCTCCAAACCAACCCAGTGTATGAGAGGCTTACTCAA	563
Oy	413	ACCCCTACTCGAACCCCTACTCTAGGTCCGTACCCAGACAGTCCGCCACACTCTCAGCTC	472
Dp	564	ATCCCTACTCTACATCTTACTACAGGCCCATACCCAGACAGCGGGCCCAACAGTACCACTG	623
Oy	473	CAAACTATCCCAACGATCTCCAGGCTCTTATATGCGCTTTGGATACCAAGATGATGAAA	532
Dp	624	CCAATACCCCAACGATTTCAAGGCGCTTGTCTGCGGCTTTGGGATATAGATGATGAAG	683
Oy	533	GCAACCAATGTGTGGATGTGGACAGAGTGTGCAACAGATTTCCACCAGTGCACACCCACC	592
Dp	684	GCAACAGATGTGTGGATGTGGACAGAGTGTGCAACAGACTCACACCAAGTGCACACCTACC	743
Oy	593	AGATCTGCATCAATACTGAAAGCGGGGTACACCTGCTCTCTGCACCGACGATATTTGGCTTC	652
Dp	744	AGATCTGTATCAACACTGAAGAGGTATACACTGCTCTCTGCACCGATGGGTACTGGCTTC	803
Oy	653	TGGAAAGGCAATGCTTTAACAATGTATGTAAATGTCCGTATGTGTACTGTCCAGACGCTCTGTG	712
Dp	804	TGGAAAGGCAATGCTTTAACAATGTATGTAAATGTCCGTATGTGTACTGTCCAGACGCTCTGTG	863
Oy	713	CGAATGTCTCGAATCTTATCTTGTATGATGCAACCTCGTGTATACCTCAATGAGATG	772
Dp	864	CAATGTCTCGAATCTTATCTTGTATGATGCAACCTCGTGTATACCTCAATGAGATG	923
Oy	773	GAAAGTCTTGCAGAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCACAACTCGC	832
Dp	924	GAAAGTCTTGCAGAGATGTGAACGAGTGTGCCAAACTGAATAACCCCTGTTCAAGACTGTG	983
Oy	833	TCAACACCTAGGGCTCTTCAATCTGCGCGCTGTGACCCAGANATGAACTTGAGAAAGATG	892
Dp	984	TCAACACCTAGGGCTCTTCAATCTGCGCGCTGTGACCCAGANATGAACTTGAGAAAGATG	1043
Oy	893	GCGTTTCAATGAGTATGTGACAGAGTGCAGCTTCTGAGTCTCTGCGCAACATGAGT	952
Dp	1044	GCAATTCAGTGAATGTATGACAGAGTGCAGCTTCTCGAAGTCTCTGTCAACACGAGT	1103
Oy	953	GTTGTGAACCAAGCCGGGACATCTTCTGCTCTCGCGCTTCAAGGCTTACATCTGTGATG	1012
Dp	1104	GTTGTGAACCAAGCCGGGCTCATTACTTCTGCTCTGCGCTTCAAGGCTTACATCTGTGATG	1163
Oy	1013	ACAACCGAAGCTGCCAAGACATCAACGAATGTGACACAGAAACACACGTCGAACCTGCG	1072
Dp	1164	ATAACCGAAGCTGCCAAGATATCAATGATGTGACACCGAAACACACGTCGAACCTGAC	1223
Oy	1073	AGCAGAGCTGCTACAATTTACAAAGGGGCTTCAAAATGCATGACACCCTACCGCTGAGG	1132
Dp	1224	TGCAGACTGCTACAATTTACAAAGGGGCTTCAAAATGATGATTCCTATCAGCTGTGAGG	1283
Oy	1133	AGCCTTATCTGAGATCAGTATAAACCGCTGTATGTGTCTGCTGAGAGAACCTGTGCTCA	1192
Dp	1284	AGCCTTATCTGAGTGTATGTGAAAAACCGCTGTATGTGTCTGCTGAGAGAACCGTGTCA	1343
Oy	1193	GAGACAGCCCTTTACCATCTTGTGACCGGGAACATGACGTGTGTCAAGACGCTCCGTTT	1252
Dp	1344	GAGACAGCCCTTTACCATCTCTGTATCGGGAACATGAGATGTGTGTCAAGACGCTCCGTTT	1403
Oy	1253	CCGCTGACATTTTCCAAATGCAAGCCACAGACCCGCTACCCGTGGGGCCATATTAACATTTCC	1312
Dp	1404	CTGCTGACATTTTCCAAATGCAAGCCACACCCGATACCCGTGGGGCCATATTAACATTTTCC	1463

Db	525	CAGATCTCCAGGCGCTTATATATGCGGCTTTGGATACGATGGATGAAGCAACCATG	584
QY	543	TGTGGATGTGGACGAGTGTGCACACAGATTCCACACAGTGCACCCCAACCAATCTGCAT	602
Db	585	TGTGGATGTGGACGAGTGTGCACACAGATTCCACACAGTGCACCCCAACCAAGATCTGCAT	644
QY	603	CAATACTGAAGCGGGGTACACTGCTCCGCGACCCGAGCGGATATTGGCTTCGTGGAAAGGCA	662
Db	645	CAATACTGAAGCGGGGTACACTGCTCCGCGACCCGAGCGGATATTGGCTTCGTGGAAAGGCA	704
QY	663	GTGCTTAGACATTGTGTAATGTGCGTATGGTTACTCCAGACACTGTGTGGAAATGTTCC	722
Db	705	GTGCTTAGACATTGTGTAATGTGCGTATGGTTACTCCAGACACTGTGTGGAAATGTTCC	764
QY	723	TGGATCTTATTTCTTGATCATGCAACCTTGTTTTACCTCTCAATGAGGATGGAAAGTCTTG	782
Db	765	TGGATCTTATTTCTTGATCATGCAACCTTGTTTTACCTCTCAATGAGGATGGAAAGTCTTG	824
QY	783	CCAAAGATGAACGAGTGTGCCACCCGAGAACCCCTGCGTGCAAACTGCGTCAACACCTA	842
Db	825	CCAAAGATGAACGAGTGTGCCACCCGAGAACCCCTGCGTGCAAACTGCGTCAACACCTA	884
QY	843	CGGCTCTTCAATCTGCGCGTGTGACCACGAGATATGAACCTTGAGGAAGATGGCGTTCAATG	902
Db	885	CGGCTCTTCAATCTGCGCGTGTGACCACGAGATATGAACCTTGAGGAAGATGGCGTTCAATG	944
QY	903	CAGTGATATGGACGAGTGTGACGCTTCTGTAGTTCCTGTGCAACATGATGTGTGAACA	962
Db	945	CAGTGATATGGACGAGTGTGACGCTTCTGTAGTTCCTGTGCAACATGATGTGTGAACA	1004
QY	963	GCCCCGACATCTCTGTCTCTGCTGCGCTTCACAGGCTACATCTCTGTGGATGACACCGAAG	1022
Db	1005	GCCCCGACATCTCTGTCTCTGCTGCGCTTCACAGGCTACATCTCTGTGGATGACACCGAAG	1064
QY	1023	CTGCCAAGACATCAACGAATGTGAGCAAGAAACCAAGCTGTGACCTGTGAGCGAAGGTG	1082
Db	1065	CTGCCAAGACATCAACGAATGTGAGCAAGAAACCAAGCTGTGACCTGTGAGCGAAGGTG	1124
QY	1083	CTACAAATTTACAAAGGGGCTTCAAAATGCATGCAGCCCATCCGCTGTGAGGAGCTTATCT	1142
Db	1125	CTACAAATTTACAAAGGGGCTTCAAAATGCATGCAGCCCATCCGCTGTGAGGAGCTTATCT	1184
QY	1143	GAGGATCAGTGATTAACCGCTGTATGTGTCTGTCTGTGAAACCTGTGTCGACAGACCAAGCC	1202
Db	1185	GAGGATCAGTGATTAACCGCTGTATGTGTCTGTCTGTGAAACCTGTGTCGACAGACCAAGCC	1244
QY	1203	CTTTACCATTGTGTAACCGGGAGACATGGAGGTGTGTACAGAGCGTCCGTTCCCGTGCAT	1262
Db	1245	CTTTACCATTGTGTAACCGGGAGACATGGAGGTGTGTACAGAGCGTCCGTTCCCGTGCAT	1304
QY	1263	CTTCCAAATGCAAAAGCCAGACCCGCTTACCTGTGGGCGTATTAATTTTCCAGATCAAAATC	1322
Db	1305	CTTCCAAATGCAAAAGCCAGACCCGCTTACCTGTGGGCGTATTAATTTTCCAGATCAAAATC	1364
QY	1323	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCT	1382
Db	1365	TGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCT	1424
QY	1383	GATGACACGGCCCATCAAAGGGGCCCGGGAAATCCAGCTGGACCTTGGAAATGATCACTGT	1442
Db	1425	GATGACACGGCCCATCAAAGGGGCCCGGGAAATCCAGCTGGACCTTGGAAATGATCACTGT	1484
QY	1443	CAACACTGTCACTCACTTCAAGAGCAGTCCGTATCCGATCCGATATATGTGTGCA	1502
Db	1485	CAACACTGTCACTCACTTCAAGAGCAGTCCGTATCCGATCCGATATATGTGTGCA	1544
QY	1503	GTAACCATTTCTGAGCTCGGGGCTTGAGGACCTCCGAGCGTCCCTCATTTGGCACCAAGGGA	1562
Db	1545	GTAACCATTTCTGAGCTCGGGGCTTGAGGACCTCCGAGCGTCCCTCATTTGGCACCAAGGGA	1604
QY	1563	CAGGAGAAAGAGGAATTAACAGAGAGATAGAGCGACACAGACGTTAGGCAATTTCTCG	1622

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REMARK	COMMENT
Db	1605	CAGAGAGAGAGAGGAATTAACAGAGAGAMTGTAGAGACGAGMACAGACTTATGACATTTCTG	1605	1605	1605	1605	1605	1605	1605	1605	1605	1605	1605
QY	1623	CTGACGTTTCCCCGGAAGAGTACGCCGACTTCTGTACTCTACCTGTACTATTG	1623	1623	1623	1623	1623	1623	1623	1623	1623	1623	1623
Db	1665	CTGACGTTTCCCCGGAAGAGTACGCCGACTTCTGTACTCTACCTGTACTATTG	1665	1665	1665	1665	1665	1665	1665	1665	1665	1665	1665
RESULT 10													
BC006636		2230 bp mRNA linear ROD 07-AUG-2002											
BC006636		mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,											
complete cds.													
BC006636		GI:13879321											
house mouse.													
mus musculus													
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.													
1 (bases 1 to 2230)													
Strausberg, R.													
Direct Submission													
Submitted (27-APR-2001) National Institutes of Health, Mammalian													
Gene Collection (MGC), Cancer Genomics Office, National Cancer													
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,													
USA													
NIH-MGC Project URL: http://mgc.ncl.nih.gov													
Contact: MGC help desk													
Email: cgabs-rt@mail.nih.gov													
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys													
CDNA Library Preparation: Life Technologies, Inc.													
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)													
DNA Sequencing by: Baylor College of Medicine Human Genome													
Sequencing Center													
Center code: BCM-HGSC													
Web site: http://www.hgsc.bcm.tmc.edu/cdna/													
Contact: amg@bcm.tmc.edu													
Gunnarite, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,													
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,													
Richards, S., Gibbs, R.A.													
Clone distribution: MGC clone distribution information can be found													
through the I.M.A.G.E. Consortium/LNL at: http://Image.lnl.gov													
Series: IRM Plate: 6 Row: f Column: 8													
This clone was selected for full length sequencing because it													
passed the following selection criteria: matched mRNA gi: 6753823.													
location/Qualifiers													
1.2230													
/organism="Mus musculus"													
/db_xref="taxon:10090"													
/map="C57BL/6J"													
/clone="MGC:5656 IMAGE:3482574"													
/tissue_type="Mammary tumor. WAP-1GF alpha model. 7 months													
old, gross tissue."													
/clone_lib="NCI_CGAP_Mam5"													
/lab_host="DH10B"													
/note="Vector: PCMV-SPORT6"													
60..1406					</								

Db 1501 TGTACTATTGACAGCTGTCACCTGCAGAGACTTGCCACCCCACTTCTATGATACAGT 1560
QY 1729 TATCAAAAAGTATTATGATGCTCCCTGATAGAAAGATGTTGGGAATTTTCAAGGCT 1788
Db 1561 TATCAAAAAGTATTATGATGCTCCCTGATAGAAAGATGTTGGGAATTTTCAAGGCT 1620
QY 1789 TCAGTTATTTCACACTATTTTCAAGAAATATAGATTAGTTGGCGGGGTCTGAGTCTAT 1848
Db 1621 TCAGTTATTTCACACTATTTTCAAGAAATATAGATTAGTTGGCGGGGTCTGAGTCTAT 1680
QY 1849 GTTCAAAAGACTGTGAACAGCTGCTGTCACTTCTTCACTTCTTCACTTCTTCACT 1908
Db 1681 GTTCAAAAGACTGTGAACAGCTGCTGTCACTTCTTCACTTCTTCACTTCTTCACT 1740
QY 1909 GTGTACTGCTTGCACAAAGACCCGGGAGCTGGCGGGGAACCTGGGAGTACTTCTTGC 1968
Db 1741 GNTTACTGCTTGCACAAAGACCCGGGAGCTGGCGGGGAACCTGGGAGTACTTCTTGC 1800
QY 1969 TTTTTCGCTACAGAGAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTTCAG 2028
Db 1801 TTTTTCGCTACAGAGAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTTCAG 1860
QY 2029 AGATGCTGTTCAAAACCATGCTGCTATTTTCAACCATTAAGAGTTTCACTTCTTCT 2088
Db 1861 AGATGCTGTTCAAAACCATGCTGCTATTTTCAACCATTAAGAGTTTCACTTCTTCT 1920
QY 2089 TAAATTGCTAATACGCTTATATCTGCTGCTGCTATTTGATTTTAAATAATATGCT 2148
Db 1921 TAAATTGCTAATACGCTTATATCTGCTGCTGCTATTTGATTTTAAATAATATGCT 1980
QY 2149 CGTAGAATTCCTTGCAGAAAGGCTTGCAGACACATGCTATG 2187
Db 1981 CGTAGAATTCCTTGCAGAAAGGCTTGCAGACACATGCTATG 2019

RESULT 8
AF093118
LOCUS Homo sapiens UP50 mRNA, complete cds.
DEFINITION AF093118
ACCESSION AF093118
VERSION AF093118.1 GI:3676821
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2019)
AUTHORS Zemel,R., Sholito,O. and Shaul,Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann
Institute of Science, Rehovot 76100, Israel
location/Qualifiers

FEATURES
source
1. 2019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue-type="urine"
1. 168
169..1515
/note="50kD; similar to human SI-5 encoded by GenBank
accession Number U03877; possible membrane and
extracellular matrix protein; urine p50 protein"
/codon_start=1
/product="up50"
/protein_id="AAC62107.1"
/db_xref="GI:3676822"
/translation="MPGIRKILVLTIALCLPSPMAOCTNFGDLDRSGCLDID
ECRTIPKACGDMACVONONGVYLCHSRTPNRYRGPSPYSGPYPAAPLASAP
NPTISRLPREGYOMDNOCYDVECATDSHOCNPTOICNNKGGYTCSTGYM
LLEGCLDIDECRYGCOOLCANVPSGSDNCGPFLINEDGRSDVDNCAETENPCV
QTCVNTYSGTRICRCDREYELBEDVHSGSDNDECSFSEFLCQHECYNOGRTFCSPPG
YILDDNRSCQDINECHRNHTCMLOQTCINLQSGFKDIDICEEPLRLSDNCKMC
PAMPGCRDQPFILYRDMDVSGRSVPADIFQOATTRYGAYVIFIKSGNEGRF
YMKQTGISATLVMTIRPIKGPRIEQLDLEMITVNTVINRGSVYIRLRIYVSOYPF"

3'UTR 1516..2019
BASE COUNT 478 a 585 c 490 g 465 t 1 others
ORIGIN
Query Match 84.0%; Score 1955.6; DB 9; Length 2019;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1998; Conservative 0; Mismatches 15; Indels 4; Gaps 3;
QY 3 CCGGGCCCTCTCCCCGCTGCTCTCTCCACAGACTCGTGGGCCCCCTGTGGAATTAACACCC 62
Db 3 CCGGGCCCTCTCCCCGCTGCTCTCTCCACAGACTCGTGGGCCCCCTGTGGAATTAACACCC 62
QY 63 GCGAGCCCCAGAGGCCAG 122
Db 63 GCGAGCCCCAGAGGCCAG 122
QY 123 CGAGCTTCTTCTGCTTGGCTTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
Db 123 CGAGCTTCTTCTGCTTGGCTTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 182
QY 183 AAGGATCTCAGCTGTTACCAATTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
Db 183 AAGGATCTCAGCTGTTACCAATTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
QY 243 ACAGTGCACGATGCTTTGACCTGATCCGACAGTCCAGAGAGAGAGAGAGAGAGAGAGAG 302
Db 243 ACAGTGCACGATGCTTTGACCTGATCCGACAGTCCAGAGAGAGAGAGAGAGAGAGAGAG 302
QY 303 ATCCCGAACCATCCCGAGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
Db 303 ATCCCGAACCATCCCGAGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
QY 363 GTATTTATGATTCCTCCCGGACAAACCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAG 422
Db 363 GTATTTATGATTCCTCCCGGACAAACCTGCTATGAGAGAGAGAGAGAGAGAGAGAGAG 422
QY 423 GACCCCTTACTCAGGTCCTGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
Db 423 GACCCCTTACTCAGGTCCTGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 483 CACGATCTCAGGCTTCTTATGCTGCTTGGATACGATGATGATGATGATGATGATGATG 542
Db 483 CACGATCTCAGGCTTCTTATGCTGCTTGGATACGATGATGATGATGATGATGATGATG 542
QY 543 TGTGATGTGAG 602
Db 543 TGTGATGTGAG 602
QY 603 CAATCTGAAAGGCGGTACACCTGCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
Db 603 CAATCTGAAAGGCGGTACACCTGCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
QY 663 GTGCTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
Db 663 GTGCTTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
QY 723 TGGATCCTATCTCTGATACATGCAACCTGCTTAACTCAATGAGATGAGATGAGATGAG 782
Db 723 TGGATCCTATCTTGTATACATGCAACCTGCTTAACTCAATGAGATGAGATGAGATGAG 782
QY 783 CCAAGATGTGAACGAGTGTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
Db 783 CCAAGATGTGAACGAGTGTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
QY 843 CGGCTCTTTCATCTGCGCTGTGTGACAGAGATATGAACTTGAGAAAGATGCGCTTCA 902
Db 843 CGGCTCTTTCATCTGCGCTGTGTGACAGAGATATGAACTTGAGAAAGATGCGCTTCA 902
QY 903 CATGTATATGAG 962
Db 903 CATGTATATGAG 962
QY 963 GCCCGGACATCTTCTGCT 1022
Db 963 GCCCGGACATCTTCTGCT 1022


```
QY 303 ATGCCAGACATCCCGAGGCGTCCGAGAGACATGATGTGTGTTAAACAAATGCGCG 362
DB 318 ATGCCAGACATCCCGAGGCGTCCGAGAGACATGATGTGTGTTAAACAAATGCGCG 377
QY 363 GATATTATGATTCCTCCCGAGCAAAACCTGTGTATCGAGGCGCTACTCGAAGCCCTACTC 422
DB 378 GATATTATGATTCCTCCCGAGCAAAACCTGTGTATCGAGGCGCTACTCGAAGCCCTACTC 437
QY 423 GACCCCTTACTAGGTCCGATACCAGAGAGTGCCTCCCACTCTCAGCTCCAACTATCC 482
DB 438 GACCCCTTACTAGGTCCGATACCAGAGAGTGCCTCCCACTCTCAGCTCCAACTATCC 497
QY 483 CACGATCTCAGGCGCTTATATATGCGCTTGTGATACAGATGGATGAAGCAACGATG 542
DB 498 CACGATCTCAGGCGCTTATATATGCGCTTGTGATACAGATGGATGAAGCAACGATG 557
QY 543 TGTGATGTGAGACGAGTGTGCAACAGATTCACACAGTGCACCCCAACCCAGATCTGCAT 602
DB 558 TGTGATGTGAGACGAGTGTGCAACAGATTCACACAGTGCACCCCAACCCAGATCTGCAT 617
QY 603 CAATACTGAAGCGGGGTACACCTGCTCCGACGAGATATTTGGCTTCTGGAAAGCCA 662
DB 618 CAATACTGAAGCGGGGTACACCTGCTCCGACGAGATATTTGGCTTCTGGAAAGCCA 677
QY 663 GTGCTTAGACATTGATGATGTGCTATGTGTTACTCCAGACAGCTCTGTGCAATGTTC 722
DB 678 GTGCTTAGACATTGATGATGTGCTATGTGTTACTCCAGACAGCTCTGTGCAATGTTC 737
QY 723 TGGATCCATTTCTTGTACATGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCTTG 782
DB 738 TGGATCCATTTCTTGTACATGCAACCCCTGTTTACCCTCAATGAGATGGAAGTCTTG 797
QY 783 CCAAGATGTGAAGAGTGTGCGCACCGAGAACCCCTGGTGCAAACTCGCTCAACACTA 842
DB 798 CCAAGATGTGAAGAGTGTGCGCACCGAGAACCCCTGGTGCAAACTCGCTCAACACTA 857
QY 843 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGATGGCTTCATTG 902
DB 858 CGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGATGGCTTCATTG 917
QY 903 CAGGATATGAGACGAGTGTGCAAGTCTCTGATGTTCTCTGCAACATGATGTGTGAACA 962
DB 918 CAGGATATGAGACGAGTGTGCAAGTCTCTGATGTTCTCTGCAACATGATGTGTGAACA 977
QY 963 GCCCGGACATCTCTGCTCTGCGCTCCAGGCTCATCTGCTGGATGAGAACAGAG 1022
DB 978 GCCCGGACATCTCTGCTCTGCGCTCCAGGCTCATCTGCTGGATGAGAACAGAG 1037
QY 1023 CTGCGAAGACATCAACGAATGTGACACAGAACACACAGTGCACACTGCGAGCAGCTG 1082
DB 1038 CTGCGAAGACATCAACGAATGTGACACAGAACACACAGTGCACACTGCGAGCAGCTG 1097
QY 1083 CTACAAATTACAAAGGGGCTTCAAAATGATGATGACCCCAACCCCTGTGAGAGAGCTTATCT 1142
DB 1098 CTACAAATTACAAAGGGGCTTCAAAATGATGATGACCCCAACCCCTGTGAGAGAGCTTATCT 1157
QY 1143 GAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGTGAGAGAGCAGCC 1202
DB 1158 GAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGTGAGAGAGCAGCC 1217
QY 1203 CTTTACCATCTTGTATACCGGGAGACATGAGAGTGTGTACAGAGCCTCGCTTCCGCTGACAT 1262
DB 1218 CTTTACCATCTTGTATACCGGGAGACATGAGAGTGTGTACAGAGCCTCGCTTCCGCTGACAT 1277
QY 1263 CTTTCAATGACAGCAGACCGCTACCCGTTGGGGCTTATTCATTTTCCAGATCAAAATC 1322
DB 1278 CTTTCAATGACAGCAGACCGCTACCCGTTGGGGCTTATTCATTTTCCAGATCAAAATC 1337
QY 1323 TGGGAATGAGGGGAGAGAAATTTACATCGGCAAAAGGGCCCATCAGTGGCCACCTGGT 1382
DB 1338 TGGGAATGAGGGGAGAGAAATTTTACATCGGCAAAAGGGCCCATCAGTGGCCACCTGGT 1397
QY 1383 GATGACACGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGACCTTGGAATGATCACTGT 1442

DB 1398 GATGACACGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGAGTGGAAATGATCACTGT 1457
QY 1443 CAACACTGTATCAACTTCAGAGAGAGCTCCGATATCCGACTGGGGATATATGTGTGCA 1502
DB 1458 CAACACTGTATCAACTTCAGAGAGAGCTCCGATATCCGACTGGGGATATATGTGTGCA 1517
QY 1503 GTACCCATCTGAGCTCGGGGAGAGAGCTCCGAGAGCTGCTCATTTGGACCAAGGGA 1562
DB 1518 GTACCCATCTGAGCTCGGGGAGAGAGCTCCGAGAGCTGCTCATTTGGACCAAGGGA 1577
QY 1563 CAGGAGAAAGAGAGAAATTAACAGAGATTAAGAGCAGACACAGCTTAGGATTTCTG 1622
DB 1578 CAGGAGAAAGAGAGAAATTAACAGAGATTAAGAGCAGACACAGCTTAGGATTTCTG 1637
QY 1623 CTGAAGCTTCCCGAAGAGTGTAGCCCGGCTTCTGACCTCACTCACTGATCTATTTGAGA 1682
DB 1638 CTGAAGCTTCCCGAAGAGTGTAGCCCGGCTTCTGACCTCACTCACTGATCTATTTGAGA 1697
QY 1683 CCTGTACCCCTGAGAGCTTGCACCCCAAGTTCTATGATAGAGTATCAAAAAGTAT 1742
DB 1698 CCTGTACCCCTGAGAGCTTGCACCCCAAGTTCTATGATAGAGTATCAAAAAGTAT 1757
QY 1743 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTCAAGGCTTCACTTATTTTCCA 1802
DB 1758 ATCATTTGCTCCCTGATAGAAAGTGTGTGTAATTTCAAGGCTTCACTTATTTTCCA 1817
QY 1803 CTATTTTCAAAAGAAATAGATTAGTTTCCGGGGGTCTGAGTATGTTCAAAAGAGCTG 1862
DB 1818 CTATTTTCAAAAGAAATAGATTAGTTTCCGGGGGTCTGAGTATGTTCAAAAGAGCTG 1877
QY 1863 AACAGCTGTCTCACTTCTTCACTCTTCACTCTCTCTCACTGTGTACTGCTTGG 1922
DB 1878 AACAGCTGTCTCACTTCTTCACTCTTCACTCTCTCTCTCACTGTGTACTGCTTGG 1937
QY 1923 CAAAGACCGGGAGGCTGGGGGAAACCCGGGAGTGTAGTGTGCTTTTGGGTACACA 1982
DB 1938 CAAAGACCGGGAGGCTGGGGGAAACCCGGGAGTGTAGTGTGCTTTTGGGTACACA 1997
QY 1983 GAGAAGGCTATGTAAACAAACCAACAGACAGATCGAAGGTTTGTAGAGATGTGTTCAA 2042
DB 1998 GAGAAGGCTATGTAAACAAACCAACAGACAGATCGAAGGTTTGTAGAGATGTGTTCAA 2057
QY 2043 AACCATGCTGTATTTTCAACCATTAAGAAAGTTCACTGTGCTTAAATTTGTATAC 2102
DB 2058 AACCATGCTGTATTTTCAACCATTAAGAAAGTTCACTGTGCTTAAATTTGTATAC 2117
QY 2103 GGTTAATTTGTCTGTGTTTCAATTTTGTAGATTTTAAAAAATGTGTAAGATTCCTTC 2162
DB 2118 GGTTAATTTGTCTGTGTTTCAATTTTGTAGATTTTAAAAAATGTGTAAGATTCCTTC 2177
QY 2163 GAAAGGCTTACAGACATGCTATGTCTGTCTCCCAACCAAGCTCTCTCATATTT 2222
DB 2178 GAAAGGCTTACAGACATGCTATGTCTGTCTCCCAACCAAGCTCTCTCATATTT 2237
QY 2223 AGCCAGAGTGTCTTGTAGAGACCCCTTAATCTGCTTCTTGTAGATTTTAAATTT 2282
DB 2238 AGCCAGAGTGTCTTGTAGAGACCCCTTAATCTGCTTCTTGTAGATTTTAAATTT 2297
QY 2283 GGATTTGAGATCAGAGTGTCTCAAACTGATTAATTAATTTGAAGAGA 2328
DB 2298 GGATTTGAGATCAGAGTGTCTCAAACTGATTAATTAATTTGAAGAAA 2343

RESULT 7
HSA133490 Homo sapiens fibulin-5. 2019 bp mRNA linear PRI 11-MAR-1999
LOCUS HSA133490
DEFINITION A133490
ACCESSION A133490.1 GI:4490529
VERSION A133490.1
KEYWORDS FTBL-5 gene; fibulin-5.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
```

```

OY 1443 CAACATGTCATCACTTCAGAGCAGCTCCGTGATCCGATCGCATATATGTGTCCA 1502
DB 1724 CAACATGTCATCACTTCAGAGCAGCTCCGTGATCCGATCGCATATATGTGTCCA 1783
OY 1503 GTACCATTTCTAGAGCTGGGCTGGAGCCTCCGACGCTGCTCATATGGCACCAGGGA 1562
DB 1784 GTACCATTTCTAGAGCTGGGCTGGAGCCTCCGACGCTGCTCATATGGCACCAGGGA 1843
OY 1563 CAGAGAGAGAGAGAAATTAACAGAGAAATGAGAGGACACAGAGCTAGGATTTCTG 1622
DB 1844 CAGAGAGAGAGAGAAATTAACAGAGAAATGAGAGGACACAGAGCTAGGATTTCTG 1903
OY 1623 CTGAAGCTTTCCCGAAGAGTACGCCGACTCTGACTCTCACTGTACTATTGCGAGA 1682
DB 1904 CTGAAGCTTTCCCGAAGAGTACGCCGACTCTGACTCTCACTGTACTATTGCGAGA 1963
OY 1683 CCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCATATGATACAGTTATCAAAAAGTAT 1742
DB 1964 CCTGTACCCCTCAGAGACTTGGCCACCCCAAGTTCATATGATACAGTTATCAAAAAGTAT 2023
OY 1743 ATCATTTGCTCCCTGTAGAGATTTGTTGTAATTTTCAAGGCTTCAGTTATTTTCCA 1802
DB 2024 ATCATTTGCTCCCTGTAGAGATTTGTTGTAATTTTCAAGGCTTCAGTTATTTTCCA 2083
OY 1803 CTATTTTCAAGAGAAATATAGATTAGTTTGGGGGGTCTGAGTCTATGTTCAAAGACTGTG 1862
DB 2084 CTATTTTCAAGAGAAATATAGATTAGTTTGGGGGGTCTGAGTCTATGTTCAAAGACTGTG 2143
OY 1863 AACAGCTTGTCTCTACCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1922
DB 2144 AACAGCTTGTCTCTACCTTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2203
OY 1923 CAAAGACCCGGAGAGCTGGGGGGGACCCCTGGAGTACCTGTTTGGTTGGTGGTACACA 1982
DB 2204 CAAAGACCCGGAGAGCTGGGGGGGACCCCTGGAGTACCTGTTTGGTTGGTGGTACACA 2263
OY 1983 GAGAAAGCTATGTAAACAAACACACAGAGATGAAAGGTTTGTAGAGATGTGTCAA 2042
DB 2264 GAGAAAGCTATGTAAACAAACACACAGAGATGAAAGGTTTGTAGAGATGTGTCAA 2323
OY 2043 AACCATGCTGTATTTTCAACCATTAAGAAGTTTCACTGTCTTAAATTTGTATAC 2102
DB 2324 AACCATGCTGTATTTTCAACCATTAAGAAGTTTCACTGTCTTAAATTTGTATAC 2383
OY 2103 GATTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2162
DB 2384 GATTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2443
OY 2163 GAAAGGCTTCAGACATGCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2222
DB 2444 GAAAGGCTTCAGACATGCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2503
OY 2223 AGCCAGTGTCTTTTGTAGAGACCCCTTAATCTTGTCTTTTGAATTTTACCCTAAT 2282
DB 2504 AGCCAGTGTCTTTTGTAGAGACCCCTTAATCTTGTCTTTTGAATTTTACCCTAAT 2563
OY 2283 GGATTTGAATTCAGAGAGCTCCCAACGATTAATATTTTGAAGAGA 2328
DB 2564 GGATTTGAATTCAGAGAGCTCCCAACGATTAATATTTTGAAGAGA 2609

```

```

RESULT 6
BC022280 2367 bp mRNA linear PRI 04-FEB-2002
LOCUS Homo sapiens, clone MGC:22412 IMAGE:4693953, mRNA, complete cds.
ACCESSION BC022280.1 GI:18490144
VERSION MGC.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2367)

```

AUTHORS

Strausberg, R.
Direct Submission
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov

COMMENT

Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Sequencing Array at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdickson@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 36 Row: n Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5453649.

FEATURES

source

1. 2367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22412 IMAGE:4693953"
/tissue_type="Lung"
/clone_id="NIH_MGC_77"
/lab_host="DH10b"
/note="Vector: pDNR-LIB"
184. 1530
/codon_start=1
/product="unknown (protein for MGC:22412)"
/protein_id="AAH22280.1"
/db_xref="GI:18490145"
/translation="MGIKRLIVTILALCLSPGNAGNOCNPFIDLRSGGCLDID
ECRTPEACGDMCVNNGCYLCIPRNPRRPSRPSYSPGSPAPAPASAP
NPTISPLICRGTQMDSENOCDVDEACATDSHOANFTQICINTEGGTCTGTG
LLEGQCLIDIDECRYGCOOLCANVPGSYSCYCNAGFTLNEDGRSCDVNECAFTENPCV
OCTVNTYGSFICRCDPELEEDVHSCDMECSFELCOHECVNPGYFSCDPG
YILDDNRSCODINCEHRNHTCMLOQCYVLMCGFKCIDIRCEPEYLRISDNCEM
PAMPGRDPTFLIYRDMDVSGRSRPAIDFQOATRYGAYYIFQIKSGNEERF
YMQGTGSIPLVMTIRPKRPRETOLDLEMTIVTFVINFROSSVIRLNIYQYF"

CDS

BASE COUNT

588 a 650 c 547 g 582 t

ORIGIN

Query Match 99.7% Score 2321.2; DB 9; Length 2367;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2323; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 3 CCGGGGCTCTCCCGGTGCTCTCAAGACTGCTGGCCCTCTGGAATAAAACACC 62
DB 18 CCGGGGCTCTCCCGGTGCTCTCAAGACTGCTGGCCCTCTGGAATAAAACACC 77
OY 63 GCGAGCCCGAGGGGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGCTCCGCCCG 122
DB 78 GCGAGCCCGAGGGGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGCTCCGCCCG 137
OY 123 CGAGCTTTCTTCGCGCTTCGATCTCTCTCGCGGCTTTGAGCATGCCAGATAAA 182
DB 138 CGAGCTTTCTTCGCGCTTCGATCTCTCTCGCGGCTTTGAGCATGCCAGATAAA 197
OY 183 AAGGATACACAGTACATTCGCTCTCTCTCTCAAGCCCTGGGAATGACAGAGC 242
DB 198 AAGGATACACAGTACATTCGCTCTCTCTCTCAAGCCCTGGGAATGACAGAGC 257
OY 243 ACAGTCACAGAAAGGCTTACCTGATTCGCAATGAGAGAGAGTATATATGAGA 302
DB 258 ACAGTCACAGAAAGGCTTACCTGATTCGCAATGAGAGAGAGTATATATGAGA 317

```


RESULT 5
LOCUS AX403659 2609 bp DNA Linear PAT 14-JUN-2002
DEFINITION Sequence 14 from Patent WO0077037.
ACCESSION AX403659
VERSION AX403659.1 GI:21437116
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1
Ashkenazi, A., Baker, K., Botstein, D., Desnovers, L., Eaton, D.L.,
Ferreira, N., Fong, W.Q., Gao, W.Q., Gerbers, H., Gierthy, M.E.,
Godard, A., Godowski, P., Gurney, A., Kijavlin, I.J., Mather, J.,
Napier, M., Pan, J., Paoni, N., Roy, M., Tamas, D., Watanabe, C.,
Williams, P.M., Wood, W.I. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0077037-A 14 21-DEC-2000;
Genentech Inc. (US)
FEATURES
source 1.2609
location/Qualifiers
BASE COUNT 622 a 718 c 622 g 647 t
ORIGIN
Query Match 99.88; Score 2322.8; DB 6; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CCGGGGCTCTCCCGGTCCTCTCCAGACTCGTCGGCCCTCTGGAAATAAACACCC 62
DB 284 CCGGGGCTCTCCCGGTCCTCTCCAGACTCGTCGGCCCTCTGGAAATAAACACCC 343
QY 63 GCGAGCCCGAGGCGCCAGAGAGGCGGAGTCCCGGAGTCCCGGGTCCCGCCG 122
DB 344 GCGAGCCCGAGGCGCCAGAGAGGCGGAGTCCCGGAGTCCCGGGTCCCGCCG 403
QY 123 CGAGCTTTCTTCGCGCTTGGCAATCTCTCGCGGCTTGGAGATGCCAGAAATAA 182
DB 404 CGAGCTTTCTTCGCGCTTGGCAATCTCTCGCGGCTTGGAGATGCCAGAAATAA 463
QY 183 AAGGATCTACTGTTACATTCCTGCTCTCTTCCAAAGCCCTGGGAATGACAGGC 242
DB 464 AAGGATCTACTGTTACATTCCTGCTCTCTTCCAAAGCCCTGGGAATGACAGGC 523
QY 243 ACAGTCTACGAATGCTTGAACCTGATTCGCGACATGACAGATGTTAGATTGATGA 302
DB 524 ACAGTCTACGAATGCTTGAACCTGATTCGCGACATGACAGATGTTAGATTGATGA 583
QY 303 ATGCCCAACCATCCCGAGGCGCTGCGAGAGAGATGATGTGTAAACCAAAATGGGG 362
DB 584 ATGCCCAACCATCCCGAGGCGCTGCGAGAGAGATGATGTGTAAACCAAAATGGGG 643

QY 363 GATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTC 422
DB 644 GATTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTC 703
QY 423 GACCCCTACTCAGTCCGATACCGAGAGTCCCGACACCTCTAGCTCCAAATATCC 482
DB 704 GACCCCTACTCAGTCCGATACCGAGAGTCCCGACACCTCTAGCTCCAAATATCC 763
QY 483 CACGATCTCAGGCTCTTATATGCGCTTTGGATACCAATGATGAAACCAATG 542
DB 764 CACGATCTCAGGCTCTTATATGCGCTTTGGATACCAATGATGAAACCAATG 823
QY 543 TGTGATGAGGAGAGTCCCAACAGATTCACACAGTGCACCCAGATCTGCAT 602
DB 824 TGTGATGAGGAGAGTCCCAACAGATTCACACAGTGCACCCAGATCTGCAT 883
QY 603 CAATCTGAAGGGGGTACATCTCTGCACCGAGAGATTTGGCTTGTGAAGGCCA 662
DB 884 CAATCTGAAGGGGGTACATCTCTGCACCGAGAGATTTGGCTTGTGAAGGCCA 943
QY 663 GTGCTTAGACATGATGATGATGCTATGCTTGTACTGCGACAGCTCTGTGGAAATGCTCC 722
DB 944 GTGCTTAGACATGATGATGATGCTATGCTTGTACTGCGACAGCTCTGTGGAAATGCTCC 1003
QY 723 TGGATCTATTTCTTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 782
DB 1004 TGGATCTATTTCTTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTG 1063
QY 783 CCAAGATGTGAACAGATGTGCCACCGAACCCTGCGTGAACCTGCTCAACACCTA 842
DB 1064 CCAAGATGTGAACAGATGTGCCACCGAACCCTGCGTGAACCTGCTCAACACCTA 1123
QY 843 CGGCTCTTCACTGCGCGTGTACCCAGATATGACTGATGAAAGATGGCGTTCATG 902
DB 1124 CGGCTCTTCACTGCGCGTGTACCCAGATATGACTGATGAAAGATGGCGTTCATG 1183
QY 903 CAGTGTATGAGAGAGTGCAGCTTCTGTAGTTCTCTGCGCAATGAGTGTGAACCA 962
DB 1184 CAGTGTATGAGAGAGTGCAGCTTCTGTAGTTCTCTGCGCAATGAGTGTGAACCA 1243
QY 963 GCCCGCACATCTTCTGCTCTGCGCTCCAGGCTTACATCTGCTGTGACCAACCGAG 1022
DB 1244 GCCCGCACATCTTCTGCTCTGCGCTCCAGGCTTACATCTGCTGTGACCAACCGAG 1303
QY 1023 CTGCCAAGACATGAAGATGTAGACAGAGAACACAGTGCAGACCTGCGAGAGCTG 1082
DB 1304 CTGCCAAGACATGAAGATGTAGACAGAGAACACAGTGCAGACCTGCGAGAGCTG 1363
QY 1083 CTACAAATTTACAGAGGGGCTTCAAAATGCATGACCCATCGCTGTGAGAGCCCTATCT 1142
DB 1364 CTACAAATTTACAGAGGGGCTTCAAAATGCATGACCCATCGCTGTGAGAGCCCTATCT 1423
QY 1143 GAGGATCAGTATGATGCTGTATGTCTCTGCTGAGAACCTGTGCTGACAGACAGCC 1202
DB 1424 GAGGATCAGTATGATGCTGTATGTCTCTGCTGAGAACCTGTGCTGACAGACAGCC 1483
QY 1203 CTTTACCATCTTATACCGGAGACATGAGAGTGTGAGAGAGCTCCGTTCCCGTGAAT 1262
DB 1484 CTTTACCATCTTATACCGGAGACATGAGAGTGTGAGAGAGCTCCGTTCCCGTGAAT 1543
QY 1263 CTTTCAAAATGCAAGCCAGCAGCCGCTACCTGCGGCTTATTTACATTTTCCAGATCAATC 1322
DB 1544 CTTTCAAAATGCAAGCCAGCAGCCGCTACCTGCGGCTTATTTACATTTTCCAGATCAATC 1603
QY 1323 TGGGAATGAGGGGAGAGAAATTTATACATGCGGCAAGGGGCCCATCACTGACCTGCTGT 1382
DB 1604 TGGGAATGAGGGGAGAGAAATTTATACATGCGGCAAGGGGCCCATCACTGACCTGCTGT 1663
QY 1383 GATGACAGCGCCCATAAAGGGGCCGCGGAAATCAGCTGACTGTGAAATGATCAGCT 1442
DB 1664 GATGACAGCGCCCATAAAGGGGCCGCGGAAATCAGCTGACTGTGAAATGATCAGCT 1723

BASE COUNT	622 a	718 c	622 g	647 t	ORIGIN
Query Match	99.88;	Score 2322.8;	DB 6;	Length 2609;	
Best Local Similarity	99.98;	Pred. No. 0;			
Matches 2324;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	3	CCCGGGCTCTCCCGCTGTCCTCTCCAGACTGCTCGGGCCCTCTGTGAATAAACACCC	62		
DB	284	CCCGGGCTCTCCCGCTGTCCTCTCCAGACTGCTCGGGCCCTCTGTGAATAAACACCC	343		
QY	63	GGGAGCCCGGAGGGCCAGAGAGAGCGAGTGCCTCCGAGCTCTCCGGGGTCCCGCG	122		
DB	344	GGGAGCCCGGAGGGCCAGAGAGAGCGAGTGCCTCCGAGCTCTCCGGGGTCCCGCG	403		
QY	123	CGAGCTTCTTCGCGCTTCGCAATCTCTCCGCGCTCTGGACATGCCAGAAATAA	182		
DB	404	CGAGCTTCTTCGCGCTTCGCAATCTCTCCGCGCTCTGGACATGCCAGAAATAA	463		
QY	183	AAGGATACACATGTTACATCTCTGCTCTCTTCCAGCCCTGGGAATGCACAGGC	242		
DB	464	AAGGATACACATGTTACATCTCTGCTCTCTTCCAGCCCTGGGAATGCACAGGC	523		
QY	243	ACAGTGCAGAAATGCTTGTACCTGTGATCCGACAGAGAGTGTATGATGATGA	302		
DB	524	ACAGTGCAGAAATGCTTGTACCTGTGATCCGACAGAGAGTGTATGATGATGA	583		
QY	303	ATGCCAATCATCTCCGAGGCTGCGAGAGAGATGATGTGTATACCAAAATGGGG	362		
DB	584	ATGCCAATCATCTCCGAGGCTGCGAGAGAGATGATGTGTATACCAAAATGGGG	643		
QY	363	GTAATTTATGATCTCCCGGACAAACCTGTGTATGAGAGGGCCCTACTCGAAACCTACTC	422		
DB	644	GTAATTTATGATCTCCCGGACAAACCTGTGTATGAGAGGGCCCTACTCGAAACCTACTC	703		
QY	423	GACCCCTTCTAGGTCCTGATCCAGAGAGAGTGCCTCCAGCTCTGAGTCCAAATATCC	482		
DB	704	GACCCCTTCTAGGTCCTGATCCAGAGAGAGTGCCTCCAGCTCTGAGTCCAAATATCC	763		
QY	483	CAGATCTCCAGGCTCTTATATGCGCTTGTGATGATGATGATGATGATGATGATG	542		
DB	764	CAGATCTCCAGGCTCTTATATGCGCTTGTGATGATGATGATGATGATGATGATG	823		
QY	543	TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	602		
DB	824	TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	883		
QY	603	CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	662		
DB	884	CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	943		
QY	663	GTCCTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	722		
DB	944	GTCCTAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1003		
QY	723	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	782		
DB	1004	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1063		
QY	783	CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	842		
DB	1064	CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1123		
QY	843	CGGCTCTTCATCTGCGCTGTGACCCAGATGATGATGATGATGATGATGATGATGATG	902		
DB	1124	CGGCTCTTCATCTGCGCTGTGACCCAGATGATGATGATGATGATGATGATGATGATG	1183		
QY	903	CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	962		
DB	1184	CAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1243		
QY	963	GGCCGGCAATATCTGCTCTGCGCTCCAGGCTATATCTGCTGATGATGATGATGATG	1022		
DB	1244	GGCCGGCAATATCTGCTCTGCGCTCCAGGCTATATCTGCTGATGATGATGATGATG	1303		
QY	1023	CTGCCAAGATATCAACGAATGTGACACAGAGACCAACAGTGTGCAACCTGACAGACG	1082		
DB	1304	CTGCCAAGATATCAACGAATGTGACACAGAGACCAACAGTGTGCAACCTGACAGACG	1363		
QY	1083	CTCAATTTCAAGAGGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG	1142		
DB	1364	CTCAATTTCAAGAGGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG	1423		
QY	1143	GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1202		
DB	1424	GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1483		
QY	1203	CTTACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1262		
DB	1484	CTTACCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1543		
QY	1263	CTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTATTTACATTTTCCAGATCAAAATC	1322		
DB	1544	CTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTATTTACATTTTCCAGATCAAAATC	1603		
QY	1323	TGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGCCCATCATGATGCCACCTGGT	1382		
DB	1604	TGGGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGCCCATCATGATGCCACCTGGT	1663		
QY	1383	GATGACACGGCCCATCAAGAGGGCCCGGGAATCCAGGTGAGCTTGGAAATGATCATCTGT	1442		
DB	1664	GATGACACGGCCCATCAAGAGGGCCCGGGAATCCAGGTGAGCTTGGAAATGATCATCTGT	1723		
QY	1443	CAACACTGTATCAACTTCAGAGGACGCTCGGTATCCGACATGCGGATATATGTGTGCA	1502		
DB	1724	CAACACTGTATCAACTTCAGAGGACGCTCGGTATCCGACATGCGGATATATGTGTGCA	1783		
QY	1503	GTAACCATTTGACACCTTCGCGGCTGAGGCTCCGAGCTGCTCTATTTGGCAACCAAGGA	1562		
DB	1784	GTAACCATTTGACACCTTCGCGGCTGAGGCTCCGAGCTGCTCTATTTGGCAACCAAGGA	1843		
QY	1563	CAGGAGAGAGAGAGAAATTAACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG	1622		
DB	1844	CAGGAGAGAGAGAGAGAAATTAACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG	1903		
QY	1623	CTGAACGTTTCCCGGAGAGAGTCCGACCTCTGACTCTACCTGATCTATTTGCA	1682		
DB	1904	CTGAACGTTTCCCGGAGAGAGTCCGACCTCTGACTCTACCTGATCTATTTGCA	1963		
QY	1683	CTGTGACCTGCGAGAGCTTCCACCCCGAGTCTATGATGATGATGATGATGATGATGATG	1742		
DB	1964	CTGTGACCTGCGAGAGCTTCCACCCCGAGTCTATGATGATGATGATGATGATGATGATG	2023		
QY	1743	ATCATTTCTCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1802		
DB	2024	ATCATTTCTCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2083		
QY	1803	CTATTTTCAAGAAATAGATTTAGTTTGGCGGGCTGTGATCTATGTTCAAAAGACTGTG	1862		
DB	2084	CTATTTTCAAGAAATAGATTTAGTTTGGCGGGCTGTGATCTATGTTCAAAAGACTGTG	2143		
QY	1863	AACAGCTGTGCTACCTTCTTCACTCTTCCACCTCTCTCTCATCTGTCTTACTGCTTGG	1922		
DB	2144	AACAGCTGTGCTACCTTCTTCACTCTTCTTCCACCTCTCTCTCATCTGTCTTACTGCTTGG	2203		
QY	1923	CAAAAGACCGGAGACTGGCGGGGAACCTGGGAGTAGTAGTTTCTTTTCCGCTGACACA	1982		
DB	2204	CAAAAGACCGGAGACTGGCGGGGAACCTGGGAGTAGTAGTTTCTTTTCCGCTGACACA	2263		
QY	1983	GAGAAGCTATGTAACAAACACAGAGATGGAAGGTTTAAAGAAATGTGTTCAA	2042		
DB	2264	GAGAAGCTATGTAACAAACACAGAGATGGAAGGTTTAAAGAAATGTGTTCAA	2323		
QY	2043	AACCAATGCTGATTTTCAACCAATTAAGAGAGTTCAGTCTCTCTTAAATTTGATTAAC	2102		
DB	2324	AACCAATGCTGATTTTCAACCAATTAAGAGAGTTCAGTCTCTCTTAAATTTGATTAAC	2383		

Db	696	CACGATCTCCAGGCGCTTTATATGCGGCTTTGGATACCAAGATGATGAAGAACCAATG	755
QY	543	TGTGATGTGGAGAGTGTGCAACAGATTCGCCACAGTGCACACCCACCAATCTGCAT	602
Db	756	TGTGATGTGGAGAGTGTGCAACAGATTCGCCACAGTGCACACCCACCAATCTGCAT	815
QY	603	CAATACTGAAGGGGGGTACACCTGCTCTGACACCGAGGATATGGCTTCTGGAAGGCA	662
Db	816	CAATACTGAAGGGGGGTACACCTGCTCTGACACCGAGGATATGGCTTCTGGAAGGCA	875
QY	663	GTCCTTGACATGTATGATGTCCTGATGTTACTGCCAGCAGCTCTGTGCAATGTTCC	722
Db	876	GTCCTTGACATGTATGATGTCCTGATGTTACTGCCAGCAGCTCTGTGCAATGTTCC	935
QY	723	TGATCCTATTCTTGTACATGACCAACCGTGTATACCCCTCAATGAGATGGAAGTCTTG	782
Db	936	TGATCCTATTCTTGTACATGACCAACCGTGTATACCCCTCAATGAGATGGAAGTCTTG	995
QY	783	CCAAGATGTGAAGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACCTA	842
Db	996	CCAAGATGTGAAGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACCTA	1055
QY	843	CGGCTCTTTCATGTCGCGCTGTGACCCAGATATGAATTTAGGAAGATGGCGTTCAATG	902
Db	1056	CGGCTCTTTCATGTCGCGCTGTGACCCAGATATGAATTTAGGAAGATGGCGTTCAATG	1115
QY	903	CAGTATATGAGAGTGTGACAGTCTCTGAGTTCTCTGCAACATGATGTTGTAACCA	962
Db	1116	CAGTATATGAGAGTGTGACAGTCTCTGAGTTCTCTGCAACATGATGTTGTAACCA	1175
QY	963	GGCGGACATACCTTCTGCTGCTGCGCTCTGAGGCTACATCTGCTGATGACCAACCGAAG	1022
Db	1176	GGCGGACATACCTTCTGCTGCTGCGCTCTGAGGCTACATCTGCTGATGACCAACCGAAG	1235
QY	1023	CTGCCAAGACATGAAGATGTGAGACAGAGAACCAACAGTGCACACCTGCGAGAGCTG	1082
Db	1236	CTGCCAAGACATGAAGATGTGAGACAGAGAACCAACAGTGCACACCTGCGAGAGCTG	1295
QY	1083	CTACATATTACAAAGGGGGCTTCAAAATGATGACACCCATCGCTGTGAGAGGCTTATCT	1142
Db	1296	CTACATATTACAAAGGGGGCTTCAAAATGATGACACCCATCGCTGTGAGAGGCTTATCT	1355
QY	1143	GAGGATCATGATTAACCCGCTGTATGTCTGCTGAGAACCCCTGCGTGCAGAGACCAACC	1202
Db	1356	GAGGATCATGATTAACCCGCTGTATGTCTGCTGAGAACCCCTGCGTGCAGAGACCAACC	1415
QY	1203	CTTACCATCTTGTACCGGAGATGAGACGTTGTAAGAGGCTCCCTTCCCGTGACAT	1262
Db	1416	CTTACCATCTTGTACCGGAGATGAGACGTTGTAAGAGGCTCCCTTCCCGTGACAT	1475
QY	1263	CTTCCAAATGCAAGCCAGACCCGCTACCCCTGGGGCTTATTACATTTCCAGATCAATC	1322
Db	1476	CTTCCAAATGCAAGCCAGACCCGCTACCCCTGGGGCTTATTACATTTCCAGATCAATC	1535
QY	1323	TGGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCATGTCACACCTGGT	1382
Db	1536	TGGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCATGTCACACCTGGT	1595
QY	1383	GATGACAGCCCATCAAAAGGGGCCCGGAAATCAAGTGGAGCTTGGAAATGATCACTGT	1442
Db	1596	GATGACAGCCCATCAAAAGGGGCCCGGAAATCAAGTGGAGCTTGGAAATGATCACTGT	1655
QY	1443	CAACACTGTATCAACTTCAGAGGAGCTCCGATCCGACTCGGATATATGTGTGCA	1502
Db	1656	CAACACTGTATCAACTTCAGAGGAGCTCCGATCCGACTCGGATATATGTGTGCA	1715
QY	1503	GTACCATTTCTGAGCTGCGGCTGTGAGCTCCGAGCTGCTCTCATTTGGCACCAGGGA	1562
Db	1716	GTACCATTTCTGAGCTGCGGCTGTGAGCTCCGAGCTGCTCTCATTTGGCACCAGGGA	1775
QY	1563	CAGAGAGAGAGAGAAATTAACAGAGATGAGAGACAGACAGCTTATGGCAATTTCTTG	1622
Db	1776	CAGAGAGAGAGAGAAATTAACAGAGATGAGAGACAGACAGCTTATGGCAATTTCTTG	1835

QY	1623	CTGAACGTTTCCCGAAGAGTCAAGCCCGAGCTTCTGACTCTACCTGTACTATTGACA	1682
Db	1836	CTGAACGTTTCCCGAAGAGTCAAGCCCGAGCTTCTGACTCTACCTGTACTATTGACA	1895
QY	1683	CGTGTACCCCTGCAAGAGCTTGCACCCCACTTCTATGATACAGTATCAAAAAGTAT	1742
Db	1896	CGTGTACCCCTGCAAGAGCTTGCACCCCACTTCTATGATACAGTATCAAAAAGTAT	1955
QY	1743	ATCATGCTCCCGATAGAAGATTTGTTGGAATTTCAAGGCGCTTATGTTATTTCA	1802
Db	1956	ATCATGCTCCCGATAGAAGATTTGTTGGAATTTCAAGGCGCTTATGTTATTTCA	2015
QY	1803	CTATTTTCAAAAAGAAATAGATTAAGTTTGGGGGGCTGAGCTATATGTTCAAGACTG	1862
Db	2016	CTATTTTCAAAAAGAAATAGATTAAGTTTGGGGGGCTGAGCTATATGTTCAAGACTG	2075
QY	1863	AACAGCTGTGTACATTTCTTCAACCTTCTTCCACCTCTTCTCTCACTGTGTACTGTTG	1922
Db	2076	AACAGCTGTGTACATTTCTTCAACCTTCTTCCACCTCTTCTCTCACTGTGTACTGTTG	2135
QY	1923	CAAAAGCCGGGAGCTGGCGGGGAACCCCTGGAGTACGTAGTTGCTTTTGGGTACCA	1982
Db	2136	CAAAAGCCGGGAGCTGGCGGGGAACCCCTGGAGTACGTAGTTGCTTTTGGGTACCA	2195
QY	1983	GAGAGGCTATGTAAACAACCAACAGAGATGAGAGGTTTGAAGATGTGTTCAA	2042
Db	2196	GAGAGGCTATGTAAACAACCAACAGAGATGAGAGGTTTGAAGATGTGTTCAA	2255
QY	2043	AACCATGCTGTGATTTTCAACCATTAAGAAGTTAGTTGTCCTTAATTTGTATAC	2102
Db	2256	AACCATGCTGTGATTTTCAACCATTAAGAAGTTAGTTGTCCTTAATTTGTATAC	2315
QY	2103	GGTTTATTTCTGCTGTTTCATTTTGAATTTTAAATAATGTGTGAATTCCTTC	2162
Db	2316	GGTTTATTTCTGCTGTTTCATTTTGAATTTTAAATAATGTGTGAATTCCTTC	2375
QY	2163	GAAAGGCTTCAAGACATGCTATGTTCTGTCTTCCAAACCAAGTCTCTCCATTTT	2222
Db	2376	GAAAGGCTTCAAGACATGCTATGTTCTGTCTTCCAAACCAAGTCTCTCCATTTT	2435
QY	2223	AGCCAGTGTCTTCTTGAAGAGCCCTTAATCTTCTTCTTGAAGATTTTACCAGAT	2282
Db	2436	AGCCAGTGTCTTCTTGAAGAGCCCTTAATCTTCTTCTTGAAGATTTTACCAGAT	2495
QY	2283	GGATTGGAATGCAAGAGTCTCCAACTGATTAATATTGAAGAGA	2328
Db	2496	GGATTGGAATGCAAGAGTCTCCAACTGATTAATATTGAAGAGA	2541

RESULT 4			
AX376340	2609 bp	DNA	linear
LOCUS			
DEFINITION	Sequence 407 from Patent WO0168848.		
ACCESSION	AX376340		
VERSION	AX376340.1		
KEYWORDS	GI:19170561		
SOURCE			
ORGANISM	human.		
REFERENCE			
AUTHORS	Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P.J., Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and Zhang, Z.		
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same		
JOURNAL	Patent: WO 0168848-A 407 20-SEP-2001;		
FEATURES			
source	Location/Qualifiers		
	1..2609		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

Db	1236	CTGCCAAGACATCAACGAATGTGAGCAGACAGAACACCAACCTGCGCAACCTGCGACGACAGCTG	1295
Oy	1083	CTACAAATTACAGGGGGCTTCAAAATGCATCGACCCATCCGCTGTGAGAGCTTATCT	1142
Db	1396	CTACAATTACAGGGGGCTTCAAAATGCATCGACCCATCCGCTGTGAGAGCTTATCT	1355
Oy	1143	GAGCATCGATATAACCGCTGTATGTGCCGTGAGAACCTGTGCGACAGACCGAC	1202
Db	1356	GAGCATCGATATAACCGCTGTATGTGCCGTGAGAACCTGTGCGACAGACCGAC	1415
Oy	1203	CTTTACCATCTTGTACCGGGACATGGAGCTGTGTCAAGAGCGTCGCTTCCGCTGACAT	1262
Db	1416	CTTTACCATCTTGTACCGGGACATGGAGCTGTGTGTCAAGAGCGTCGCTTCCGCTGACAT	1475
Oy	1263	CTTCCAAATGCAACCCACGACCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATC	1322
Db	1476	CTTCCAAATGCAACCCACGACCCGCTACCTGTGGGGCTATTACATTTTCCAGATCAATC	1535
Oy	1323	TGGGAATGAGGGGACAGAAATTTATCATGCGGCAAAAGGGGCCCATACGTGCACCTGTG	1382
Db	1536	TGGGAATGAGGGGACAGAAATTTATCATGCGGCAAAAGGGGCCCATACGTGCACCTGTG	1595
Oy	1383	GATGACACGGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGGACTTGGAAATGATCTGT	1442
Db	1596	GATGACACGGCCCATCAAAAGGGCCCCGGGAAATCCAGCTGGACTTGGAAATGATCTGT	1655
Oy	1443	CACACAGTCATCAACTCAGAGGACAGCTCCGTATCCGACTTGGGGATATATGTGTGCA	1502
Db	1656	CACACAGTCATCAACTCAGAGGACAGCTCCGTATCCGACTTGGGGATATATGTGTGCA	1715
Oy	1503	GTACCCATTCTGAGCCCTCGGGCTGGAGCCTCCAGCGCTCCCTCATTTGGCACCAGAGGA	1562
Db	1716	GTACCCATTCTGAGCCCTCGGGCTGGAGCCTCCAGCGCTCCCTCATTTGGCACCAGAGGA	1775
Oy	1563	CAGAGAAAGAGAAATTAACAGAGAGATAGAGCGACACACAGAGCTTAGGCATTTCTG	1622
Db	1776	CAGAGAAAGAGAAATTAACAGAGAGATAGAGCGACACACAGAGCTTAGGCATTTCTG	1835
Oy	1623	CTGAAGCTTTCCCGAAGAGTCAACCCCGACTTCTGTCACTCACTGATATATGAGAGA	1682
Db	1836	CTGAAGCTTTCCCGAAGAGTCAACCCCGACTTCTGTCACTCACTGATATATGAGAGA	1895
Oy	1683	CTGTGCACCCTCGACGAGACTTCCACCCCCCACTTCCATGATACGATTATCAAAAAGTAT	1742
Db	1896	CTGTGCACCCTCGACGAGACTTCCACCCCCCACTTCCATGATACGATTATCAAAAAGTAT	1955
Oy	1743	ATCAATGCTCCCGATGATGAGATTTGTGGGAATTTCAAGGGCTTCAGTTATTTCCA	1802
Db	1956	ATCAATGCTCCCGATGATGAGATTTGTGGGAATTTCAAGGGCTTCAGTTATTTCCA	2015
Oy	1803	CTATTTTCAAGAAATATAGATTAGATTGCGGGGCTGAGCTCTATGTTCAAGACGTG	1862
Db	2016	CTATTTTCAAGAAATATAGATTAGATTGCGGGGCTGAGCTCTATGTTCAAGACGTG	2075
Oy	1863	AACGAGTTGCTGTACTCTTCCACTCTTCCACTCTTCTCACTGTGTACTGTG	1922
Db	2076	AACGAGTTGCTGTACTCTTCCACTCTTCCACTCTTCTCACTGTGTACTGTG	2135
Oy	1923	CAAGACCCGGGAGCTGGCGGGGAACCTGGAGATAGCTATGTTTGTGCTTACACA	1982
Db	2136	CAAGACCCGGGAGCTGGCGGGGAACCTGGAGATAGCTATGTTTGTGCTTACACA	2195
Oy	1983	GAGAAAGCTATGTAACAAACACACAGCATGCAAGGGTTTGAAGATGTGTTCAA	2042
Db	2196	GAGAAAGCTATGTAACAAACACACAGCATGCAAGGGTTTGAAGATGTGTTCAA	2255
Oy	2043	AACCATGCTGTATTTTCAACCAATTAAGAAAGTTTCAAGTGTGCTTAATTTGTATAC	2102
Db	2256	AACCATGCTGTATTTTCAACCAATTAAGAAAGTTTCAAGTGTGCTTAATTTGTATAC	2315
Oy	2103	GATTAAATCTGTCTGTCAATTTGATATTTTAAAAAATATGTGTGAATCTCTTC	2162

Dd		2316	GGTTTAATTCGTCTGTTTCAATTTTGAGATATTTTTTAAAAAANFNTGCTACAAATTCCTC	2375
Qy		2163	GAAAGCCCTTAGACACATCGTATGTTCTGTCTCTCCAACACCAGTCTCTCCATTTT	2222
Dd		2376	GAAAGGCCCTTAGACACATCGTATGTTCTGTCTCTCCAACACCAGTCTCTCCATTTT	2435
Qy		2223	AGCCGAGTGTCTTTCTTTAGAGCCCGCTTAATCTGCTCTTTGAATTTTACCACATT	2282
Dd		2436	AGCCGAGTGTCTTTCTTTAGAGCCCGCTTAATCTGCTCTTTGAATTTTACCACATT	2495
Qy		2283	GGATTGGAATGCAGAGGTCCTCAAACACTGATTAATTTTGAAGA	2328
Dd		2496	GGATTGGAATGCAGAGGTCCTCAAACACTGATTAATTTTGAAGA	2541
<hr/>				
RESULT 3				
LOCUS AR173204 2550 bp DNA linear PAT 17-DEC-2001				
DEFINITION Sequence 2 from patent US 6303765.				
ACCESSION AR173204				
VERSION AR173204.1 GI:17912695				
KEYWORDS				
SOURCE Unknown.				
ORGANISM Unclassified.				
REFERENCE 1 (bases 1 to 2550)				
AUTHORS Bandman,O., Corley,N.C. and Guegler,K.J.				
TITLE Human extracellular matrix proteins				
JOURNAL Patent: US 6303765-A 2-16-OCT-2001;				
FEATURES Location/Qualifiers				
source 1..2550				
BASE COUNT 623 a 690 c 596 g 641 t				
ORIGIN /organism="unknown"				
<hr/>				
Query Match 99.8%; Score 2324.4; DB 6; Length 2550;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 2325; Conservativeness 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	3	CCCGGCGCTCCCGGCTCTCTCTCCAGACATCGCTCGGCCCTCTGGAATAAAAACACC	62	
Dd	216	CCGGGCGCTCTCCCGGTCTCTCTCCAGACATCGCTCGGCCCTCTGGAATAAAAACACC	275	
Qy	63	GCGAGCCCCGAGGGCCGAGAGGCGGAGCTGCCGAGCTCTCCGGGGTCCCGCCG	122	
Dd	276	GCGAGCCCCGAGGGCCGAGAGGCGGAGCTGCCGAGCTCTCCGGGGTCCCGCCG	335	
Qy	123	CGAGGTTCTCTCGGCGCTTGCATCTCTCTCCGCGGGCTTTGGACATGCCAGGAATAA	182	
Dd	336	CGAGGTTCTCTCGGCGCTTGCATCTCTCTCCGCGGGCTTTGGACATGCCAGGAATAA	395	
Qy	183	AAGATATCTACTGTTTACCATTTGSGCTCTCTGTCTTCCAAGCCCTGGGAATGACAAGC	242	
Dd	396	AAGATATCTACTGTTTACCATTTGSGCTCTCTGTCTTCCAAGCCCTGGGAATGACAAGC	455	
Qy	243	ACAGTGCACGAATGGCTTTGACCTTGGATCGCCAGTCAGAGACATGTTAGATTATGATGA	302	
Dd	456	ACAGTGCACGAATGGCTTTGACCTTGGATCGCCAGTCAGAGACATGTTAGATTATGATGA	515	
Qy	303	ATGCGGAACATCCCGCGAGGCTCGCGGAGAGACATGATGTGTAAACAAAATGGCGG	362	
Dd	516	ATGCGGAACATCCCGCGAGGCTCGCGGAGAGACATGATGTGTAAACAAAATGGCGG	575	
Qy	363	GTATTATGATTCCTCCGAGACAAACCTGTGTATGAGAGGGCTACTTGAACCCCTACTC	422	
Dd	576	GTATTATGATTCCTCCGAGACAAACCTGTGTATGAGAGGGCTACTTGAACCCCTACTC	635	
Qy	423	GACCCCTACACAGGTCCGTACCCAGACAGCTGCCCCACATCTCACAGTCCAAACTATCC	482	
Dd	636	GACCCCTACACAGGTCCGTACCCAGACAGCTGCCCCACATCTCACAGTCCAAACTATCC	695	
Qy	483	CACATCTCCACGGGCTCTTATATGACCGTTTGGATTACAGATGGATGAAGAACACATATG	542	

```

QY 1561 GACAGAGAGAGAGAGAAATTAACAGAGAGATAGAGACGACAGACGATTTAGCATTTCC 1620
DB 1561 GACAGAGAGAGAGAGAGAAATTAACAGAGAGATAGAGACGACAGACGATTTAGCATTTCC 1620
QY 1621 TGCTAAGCTTTTCCCGAAGAGTACGCCCGACTTCTGACTCTACCTGTACTATTGCA 1680
DB 1621 TGCTAAGCTTTTCCCGAAGAGTACGCCCGACTTCTGACTCTACCTGTACTATTGCA 1680
QY 1681 GACCTGTACCCCTGGAGAGTACGCCCGACTTCTGACTCTACCTGTACTATTGCA 1740
DB 1681 GACCTGTACCCCTGGAGAGTACGCCCGACTTCTGACTCTACCTGTACTATTGCA 1740
QY 1741 TTATCATTTGCTCCCTGTATAGAGATTTGTGTAATTTTCAAGGCGCTTCAGTTATTTTC 1800
DB 1741 TTATCATTTGCTCCCTGTATAGAGATTTGTGTAATTTTCAAGGCGCTTCAGTTATTTTC 1800
QY 1801 CACTATTTTCAAGAAATAGATTTAGTTTGGGGGTCTGAGTCTATGTTCAAGACTG 1860
DB 1801 CACTATTTTCAAGAAATAGATTTAGTTTGGGGGTCTGAGTCTATGTTCAAGACTG 1860
QY 1861 TGAACAGCTTGTCTCACTCTTCACTCTTCCACTCTCTCTCACTGTGTTTCTGCT 1920
DB 1861 TGAACAGCTTGTCTCACTCTTCACTCTTCCACTCTCTCTCACTGTGTTTCTGCT 1920
QY 1921 TGCAAGAGCCCGGAGCTGCGGGGAGCCCTGGAGTAGCTAGTTTCTTTTTCGTA 1980
DB 1921 TGCAAGAGCCCGGAGCTGCGGGGAGCCCTGGAGTAGCTAGTTTCTTTTTCGTA 1980
QY 1981 CAGAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTATAGACATGTGTTTC 2040
DB 1981 CAGAGAGGCTATGTAAACAAACACAGAGATCGAAGGTTTATAGACATGTGTTTC 2040
QY 2041 AAAACCATGCTGATTTTCAACATTAAGAAAGTTTCACTGTCTTAAATTTGATA 2100
DB 2041 AAAACCATGCTGATTTTCAACATTAAGAAAGTTTCACTGTCTTAAATTTGATA 2100
QY 2101 ACGGTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2101 ACGGTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
QY 2161 TCGAAGAGGCTTCAACACATGCTATGTTCTGCTTCCCAACCCAGTCTCTCCAT 2220
DB 2161 TCGAAGAGGCTTCAACACATGCTATGTTCTGCTTCCCAACCCAGTCTCTCCAT 2220
QY 2221 TTAGCCAGTGTCTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCCA 2280
DB 2221 TTAGCCAGTGTCTTCTTGGAGACCCCTTAATCTGCTTCTTGAATTTTACCCA 2280
QY 2281 TTGATTTGGAATGACAGAGTCTCAAACTGATTTAAATTTTGAAGAGA 2328
DB 2281 TTGATTTGGAATGACAGAGTCTCAAACTGATTTAAATTTTGAAGAGA 2328

```

```

RESULT 2
AR036548 2550 bp DNA linear PAT 29-SEP-1999
LOCUS AR036548
DEFINITION Sequence 2 from patent US 5872234.
ACCESSION AR036548
VERSION AR036548.1 GI:5953216
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2550)
AUTHORS Bandman,O., Corley,N.C. and Guejler,K.J.
TITLE Human extracellular matrix proteins
JOURNAL Patent: US 5872234-A 2 16-FEB-1999;
FEATURES
source location/Qualifiers
BASE COUNT 623 a 690 c 596 g 641 t
ORIGIN

```

```

Query Match 99.8%; Score 2324.4; DB 6; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 CCCGGGCTCTCCCGCTGCTCTCTCCAGACTGCTGGGCGCTCTGGAATTAACACACC 62
DB 216 CCCGGGCTCTCCCGCTGCTCTCTCCAGACTGCTGGGCGCTCTGGAATTAACACACC 275
QY 63 GCGAGCCCGGAGGGCCCAAGAGAGCCGACGCTGCCAGCTCTCCGGGGTCCCGCCG 122
DB 276 GCGAGCCCGGAGGGCCCAAGAGAGCCGACGCTGCCAGCTCTCCGGGGTCCCGCCG 335
QY 123 CGAGCTTCTTCCGCTTGGCATCTCTCTGCGGGGTCTTGAACATGCCAGAAATAA 182
DB 336 CGAGCTTCTTCCGCTTGGCATCTCTCTGCGGGGTCTTGAACATGCCAGAAATAA 395
QY 183 AAGGATCTCACTGTTACCATCTGCTCTGCTCTTCAAGCCCTGGGAATGACAGGC 242
DB 396 AAGGATCTCACTGTTACCATCTGCTCTGCTCTTCAAGCCCTGGGAATGACAGGC 455
QY 243 AAGTGCACGAATGGCTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
DB 456 AAGTGCACGAATGGCTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
QY 303 ATGCCGAACATCCCGAGGCTGCGAGAGACATATGTTTAAACCAAAATGGCGG 362
DB 516 ATGCCGAACATCCCGAGGCTGCGAGAGACATATGTTTAAACCAAAATGGCGG 575
QY 363 GTATTTATCATTTCCCGGACAAACCTGTGTATGAGAGGCTTCTGAAACCCCTACTC 422
DB 576 GTATTTATCATTTCCCGGACAAACCTGTGTATGAGAGGCTTCTGAAACCCCTACTC 635
QY 423 GACCCCTCTCAAGTCCGATCCGACAGAGTCCGACCTGCTGCTGCTGCTGCTGCT 482
DB 636 GACCCCTCTCAAGTCCGATCCGACAGAGTCCGACCTGCTGCTGCTGCTGCTGCT 695
QY 483 CAGATCTCAGGCTCTTATATGCGCTTGGATATACATGATGATGATGATGATGATG 542
DB 696 CAGATCTCAGGCTCTTATATGCGCTTGGATATACATGATGATGATGATGATGATG 755
QY 543 TGTGATGTGAGAGAGTGTGCAACAGATTTCCACACAGTGCACCCACCCAGATCTGCAT 602
DB 756 TGTGATGTGAGAGAGTGTGCAACAGATTTCCACACAGTGCACCCACCCAGATCTGCAT 815
QY 603 CAATCTGAAAGGCGGTACACCTGCTCTGACACAGAGATTTGGCTTGGAAAGGCA 662
DB 816 CAATCTGAAAGGCGGTACACCTGCTCTGACACAGAGATTTGGCTTGGAAAGGCA 875
QY 663 GTGCTTAGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
DB 876 GTGCTTAGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 935
QY 723 TGGATCTATTTCTTGTACATGCAACCTGTTTACCCTCAATGAGATGGAAGTCTTG 782
DB 936 TGGATCTATTTCTTGTACATGCAACCTGTTTACCCTCAATGAGATGGAAGTCTTG 995
QY 783 CCAAGATGGAAGAGTGTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
DB 996 CCAAGATGGAAGAGTGTGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1055
QY 843 CGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
DB 1056 CGGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
QY 903 CAGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962
DB 1116 CAGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1175
QY 963 GCGCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
DB 1176 GCGCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1235
QY 1023 CTGCCAAGACATCAAGATGTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082

```

TITLE Matsumori, A., Sasayama, S., Chien, K. R., and Honjo, T.
DANCE, a novel secreted RGD protein expressed in developing,
atherosclerotic, and balloon-injured arteries
J. Biol. Chem. 274 (32), 22476-22483 (1999)
MEDLINE 99357779
PUBMED 10428823
2 (bases 1 to 2328)
REFERENCE Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
TITLE Direct Submission
Submitted (09-DEC-1998) Medicine, University of California San
Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
JOURNAL location/Qualifiers
1..2328
FEATURES
source

CDS

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q32.1"
169..1515
/note="DANCE: contains 6 cbpRF domains and 1 RGD motif;
secreted protein; expressed in developing aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="AA041768.1"
/db_xref="GI:5305673"
/translation="MPGIRKRLITVITLALCLPSRPMQAQCTNGEDLRSGCCLDID
ECTPIEPACRDMVCMVNGSYLCIPRNPYRGPYSPYSPYPAAPPLASAP
NPTISPLRFGYOMDSQVDEACSHQCNPTQICINTEGVCCTGDMY
LLEGOCIDIDECRGGYCOQLCANPAGSYSCNPGFTLNEGSCODVACATENPCV
QTCVNTGSEFCRCDPELEEDENGVHSCSDMECSFEFLQHECVNPOPTVYCSGPG
YLLDNRSCODINECHRNHTCNLQOCVYVLOGGFCIDIPRCEPVLRIISDNRCM
PAMPGCRDPTFLYEDMDVYSGRSVPADIFQMAATRYRGAATYIFQISNEGRF
YMQTGISALVMTNPRIKGRPREQLDLEMTVMVINFRSSVRLRLVSYPR"
BASE COUNT 560 a 646 c 541 g 581 t
ORIGIN

Query Match 100.0%; Score 2328; DB 9; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACCCGGCGCTCCCGGTCTCTCCACGACTGCTGGCCCTCTGGAATAAAACAC 60.
DB 1 GACCCGGCGCTCCCGGTCTCTCCACGACTGCTGGCCCTCTGGAATAAAACAC 60
QY 61 CCGCGAG 120
DB 61 CCGCGAG 120
QY 121 CCGGAGCTTCTCTGCGCTCGCATCTCTCTGCGCGCTTGGACATGCCAGAGATA 180
DB 121 CCGGAGCTTCTCTGCGCTCGCATCTCTCTGCGCGCTTGGACATGCCAGAGATA 180
QY 181 AAAAGATCTACTGTTACATCTGGCTCTGCTTCCAAGCCCTGGGAATGACAG 240
DB 181 AAAAGATCTACTGTTACATCTGGCTCTGCTTCCAAGCCCTGGGAATGACAG 240
QY 241 GCACAGAGCAGATGGCTTGGACCTGGATGCCAGTACAGACATGTTAATATATGAT 300
DB 241 GCACAGAGCAGATGGCTTGGACCTGGATGCCAGTACAGACATGTTAATATATGAT 300
QY 301 GAATGCGAAGCATCCCGAGAGCTGCGGAGAGACATGATGTGTTAAACAAAATGGC 360
DB 301 GAATGCGAAGCATCCCGAGAGCTGCGGAGAGACATGATGTGTTAAACAAAATGGC 360
QY 361 GGGTATTATGATCTCCCGGACAAACCTGTGATGAGAGGCGCTTACTGAAACCCCTAC 420
DB 361 GGGTATTATGATCTCCCGGACAAACCTGTGATGAGAGGCGCTTACTGAAACCCCTAC 420
QY 421 TCGACCCCTCTACTAGTCCGTACCCACAGACCTGCCACCACTCTCAACTATAT 480
DB 421 TCGACCCCTCTACTAGTCCGTACCCACAGACCTGCCACCACTCTCAACTATAT 480

QY 481 CCCACGATCTCCAGAGCCCTTATATAGCCGCTTGGATACAGATGGATGAAGAACACAA 540
DB 481 CCCACGATCTCCAGAGCCCTTATATAGCCGCTTGGATACAGATGGATGAAGAACACAA 540
QY 541 TGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGATGACACCCACAGATCTGC 600
DB 541 TGTGTGATGTGAGAGAGTGTGCAACAGATTCCACAGATGACACCCACAGATCTGC 600
QY 601 ATCAATATCTGAAGCGGGGTACACCTGCTCTGACACGAGATATGGCTTGTGAAGGC 660
DB 601 ATCAATATCTGAAGCGGGGTACACCTGCTCTGACACGAGATATGGCTTGTGAAGGC 660
QY 661 CAGTCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 CAGTCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 CCTGATCTCTATTTCTTGTACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCT 780
DB 721 CCTGATCTCTATTTCTTGTACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCT 780
QY 781 TGCCAAAGATGTGAAGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACC 840
DB 781 TGCCAAAGATGTGAAGAGTGTGCCACCGAGAACCCCTGCTGCAAACTGCTCAACACC 840
QY 841 TAGGGCTCTTCAATCTGCGCGCTGTGACCCAGATATGAACTTGAGAGATGGCTTCAT 900
DB 841 TAGGGCTCTTCAATCTGCGCGCTGTGACCCAGATATGAACTTGAGAGATGGCTTCAT 900
QY 901 TGCAGATATATGAGAGAGTGTGAGCTTCTGTGATTCCTGTGCAACATGATGTGTGAA 960
DB 901 TGCAGATATATGAGAGAGTGTGAGCTTCTGTGATTCCTGTGCAACATGATGTGTGAA 960
QY 961 CAGCCCGGACATATCTTGTCTCTGCGCCCTCCAGGCTATATCTGTGATGACAAACCA 1020
DB 961 CAGCCCGGACATATCTTGTCTCTGCGCCCTCCAGGCTATATCTGTGATGACAAACCA 1020
QY 1021 AGCTGCCAAGACATCAACGAATGTGAGACAGAACCAACAGCTGCAACCTGACAGACAG 1080
DB 1021 AGCTGCCAAGACATCAACGAATGTGAGACAGAACCAACAGCTGCAACCTGACAGACAG 1080
QY 1081 TGCTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCGCTGTAGAGAGCTTAT 1140
DB 1081 TGCTACAAATTTACAAAGGGGCTTCAAAATGATGACACCCATCGCTGTAGAGAGCTTAT 1140
QY 1141 CTGAGATACAGTGAATACCGCTGTATGTCTGTGAGAACCTGCGTGCAGAGACAG 1200
DB 1141 CTGAGATACAGTGAATACCGCTGTATGTCTGTGAGAACCTGCGTGCAGAGACAG 1200
QY 1201 CCCCTTACCATCTTGTACCGGAGACATGAGAGCTGTGTCAGACGCTCCGCTCCGCTGAC 1260
DB 1201 CCCCTTACCATCTTGTACCGGAGACATGAGAGCTGTGTCAGACGCTCCGCTCCGCTGAC 1260
QY 1261 ATCTTCCAAATGACAGACAGACAGACCGCTTACCTGTGGGCTATATATTTCCAGATCAA 1320
DB 1261 ATCTTCCAAATGACAGACAGACAGACCGCTTACCTGTGGGCTATATATTTCCAGATCAA 1320
QY 1321 TCTGGAATGAGGGGAGAGAAATTTACATGAGGGGCAAGGGGCCCATGATGACACCTG 1380
DB 1321 TCTGGAATGAGGGGAGAGAAATTTACATGAGGGGCAAGGGGCCCATGATGACACCTG 1380
QY 1381 GTGATGACACGCCCATCAAAAGGGGCCGGGGAATTCACACTGACTTGGAAATGATCAGT 1440
DB 1381 GTGATGACACGCCCATCAAAAGGGGCCGGGGAATTCACACTGACTTGGAAATGATCAGT 1440
QY 1441 GTCAACACTGTATCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 GTCAACACTGTATCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 CAGTACCATCTGTAGGCTCGGGGCTGAGAGCTGCGAGCTGTGCTCATTTGGACCAAGG 1560
DB 1501 CAGTACCATCTGTAGGCTCGGGGCTGAGAGCTGCGAGCTGTGCTCATTTGGACCAAGG 1560

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 6070.42 Seconds

(Without alignments)
11160.894 Million cell updates/sec

Title: US-09-674-379A-12

Perfect score: 2328
Sequence: 1 gaccggcgctctcccgctg.....tgattaatatggaaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: 1: gb_ba: 2: gb_hlg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_ph: 7: gb_pl: 8: gb_pl: 9: gb_pl: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_hlg_hum: 31: em_hlg_inv: 32: em_hlg_other: 33: em_hlg_mus: 34: em_hlg_pin: 35: em_hlg_rtd: 36: em_hlg_mam: 37: em_hlg_vrt: 38: em_sy: 39: em_hlg_hum: 40: em_hlg_mus: 41: em_hlg_other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2328	100.0	2328	9 AF112152	AF112152 Homo sapi
2	2324.4	99.8	2550	6 AR036548	AR036548 Sequence
3	2324.4	99.8	2550	6 AR173204	AR173204 Sequence
4	2322.8	99.8	2609	6 AX376340	AX376340 Sequence
5	2322.8	99.8	2609	6 AX403659	AX403659 Sequence
6	2321.2	99.7	2367	9 BC022280	BC022280 Homo sapi
7	2014.8	86.5	2019	9 HSA13490	HSA13490 Homo sapi
8	1955.6	84.0	2019	9 AF093118	AF093118 Homo sapi
9	1674.2	71.9	1720	6 A84086	A84086 Sequence 1
10	1481.8	63.7	2230	10 BC006636	BC006636 Mus muscu
11	1481.8	63.7	2478	10 AF112151	AF112151 Mus muscu
12	1476.6	63.4	2499	10 AF112153	AF112153 Rattus no
13	1475.2	63.4	2304	10 AF137350	AF137350 Rattus no
14	974.4	41.9	171901	9 AC007057	AC007057 Homo sapi
15	974.4	41.9	175210	9 CNS07EG5	AL590338 Human chr
16	973.4	41.8	2324	9 AK094032	AK094032 Homo sapi
17	630.4	27.1	648	9 HSM800793	AL110148 Homo sapi
18	501	21.5	172341	2 AC095490	AC095490 Rattus no
19	387.6	16.6	1440	9 HSA132819	AJ132819 Homo sapi
20	387.6	16.6	1536	9 AB030655	AB030655 Homo sapi
21	386	16.6	1480	6 AX023967	AX023967 Sequence
22	386	16.6	1561	9 BC010456	BC010456 Homo sapi
23	386	16.6	1707	9 AK000980	AK000980 Homo sapi
24	386	16.6	1875	6 AX201326	AX201326 Sequence
25	386	16.6	1875	6 AX464184	AX464184 Sequence
26	386	16.6	2018	6 AR095382	AR095382 Sequence
27	384.4	16.5	1757	9 AF109121	AF109121 Homo sapi
28	379.6	16.3	1825	9 AF124486	AF124486 Homo sapi
29	378.6	16.3	1705	9 AF093119	AF093119 Homo sapi
30	353.4	15.2	1778	10 AF046870	AF046870 Cricethlu
31	352.4	15.1	1358	6 AX023976	AX023976 Sequence
32	350.2	15.0	1512	10 AF104223	AF104223 Mus muscu
33	350.2	15.0	1513	6 AX023961	AX023961 Sequence
34	350.2	15.0	1794	10 AF109122	AF109122 Mus muscu
35	348.6	15.0	1724	10 BC012269	BC012269 Mus muscu
36	348.2	15.0	518	6 AX079354	AX079354 Sequence
37	321.8	13.8	340	6 AX332820	AX332820 Sequence
38	321.8	13.8	340	6 AX334715	AX334715 Sequence
39	321.8	13.8	340	6 AX336178	AX336178 Sequence
40	315	13.5	1021	6 AX023954	AX023954 Sequence
41	295.8	12.7	2613	9 AK098186	AK098186 Homo sapi
42	268	11.5	1132	6 AX023965	AX023965 Sequence
43	260.6	11.2	172674	9 CDS0000R	AL049872 Human chr
44	251	10.8	1677	10 BC031184	BC031184 Mus muscu
45	250.6	10.8	1911	10 D89730	D89730 Rattus ratt

ALIGNMENTS

RESULT 1
AF112152
LOCUS
DEFINITION Homo sapiens developmental arteries and neural crest EGF-like
protein mRNA, complete cds.
ACCESSION AF112152
VERSION AF112152
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2328)
Nakamura,T., Ruiz-Lozano,P., Lindner,V., Yabe,D., Taniwaki,M.,
Furukawa,Y., Kobuke,K., Tashiro,K., Lu,Z., Andon,N.L., Schaub,R.,

Not1; Site 2; Sal1; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGGACCCAGCGTCCG-3' and
5'-GACTATCTTCAGATCCGAGAGCGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 225 a 284 c 245 g 179 t
ORIGIN

Query Match 42.9%; Score 576.4; DB 14; Length 933;
Best Local Similarity 94.6%; Pred. No. 9e-149;
Matches 619; Conservative 0; Mismatches 31; Indels 4; Gaps 2;

```
QY 602 ATGGAAGCTCTTGCCCAAGATGTAGACGATGTGCGCCAGAGACCCCTGCGCAACCT 661
DB 2 ATGGAAGGCTCTTCCCAAGATGTAGACGATGTGCGCCAGAGACCCCTGCGCAACCT 61
QY 662 GCGTCAACACCTACCGCTCTTTCATCTGGCGGCTGTGACCCAGATATGAATGAGGAG 721
DB 62 GCGTCAACACCTACCGCTCTTTCATCTGGCGGCTGTGACCCAGATATGAATGAGGAG 121
QY 722 ATGGCGTTCATTGCAATGATATGAGAGAGTGCAGCTTCTGAGTTCCTCTGCCAATG 781
DB 122 ATGGCGTTCATTGCAATGATATGAGAGAGTGCAGCTTCTGAGTTCCTCTGCCAATG 181
QY 782 AGTGTGTGAACCAAGCCCGGCACATCTTCTGCTCTGCTCCCTCCAGGCTACATCTGTGG 841
DB 182 AGTGTGTGAACCAAGCCCGGCACATCTTCTGCTCTGCTCCCTCCAGGCTACATCTGTGG 241
QY 842 ATGACACGAGAGCTGCGCAAGATCAAGAAATGTGAGACAGAGAACCCAGCTGCAAC 901
DB 242 ATGACACGAGAGCTGCGCAAGATCAAGAAATGTGAGACAGAGAACCCAGCTGCAAC 301
QY 902 TGCAGAGAGAGCTGCTCAATTTTACAAGGGGGCTTCAATGCATGACCCCAATCCGCTGTG 961
DB 302 TGCAGAGAGAGCTGCTCAATTTTACAAGGGGGCTTCAATGCATGACCCCAATCCGCTGTG 361
QY 962 AGGAGCCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCCCTGGCT 1021
DB 362 AGGAGCCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCCCTGGCT 421
QY 1022 GCAGAGACAGCGCCTTACCATCTTGTATCGGGGACATGAGACGTGTGCAGAGAGCTCCG 1081
DB 422 GCAGAGACAGCGCCTTACCATCTTGTATCGGGGACATGAGAGGTGTGCAGAGAGCTCCG 481
QY 1082 TTCCCGCTGACATCTTCAAAATGCAAGCCAGACCCGCTACCTGCGGCTATTACATTT 1141
DB 482 TTCCCGCTGACATCTTCAAAATGCAAGCCAGACCCGCTACCTGCGGCTATTACATTT 541
QY 1142 TCCAGATCAAAATCTGGGAATGA-GGGCAGAGAAATTTTACATGCGGCAAA---CGGGCCCC 1197
DB 542 TCCAGATCAAAATCTGGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAACGCGGCCCC 601
QY 1198 ATCAGTCCAGCCCTGTGATGACACGCGCATCAAGGGCCCCGGGAATCCAG 1251
DB 602 ATCAGTCCAGCCCTGTGATGAGACCGCGCATTCCTCAAGGGGGCCCCGG 655
```

Search completed: July 3, 2003, 17:36:03
Job time : 1988.41 secs

QY 866 TCAGCAATGTGACGACAGGACCAACGTCGACACTGACAGCAGCGTGTACATTTAC 925
 DB 541 TCATTAATATGTGACGACCAACGACGATGTACCTGACGACAGCTTGCACATCTAC 600
 QY 926 AAGGGGGCTTCAATATGATGACGACGACGACGACGACGACGACGACGACGACGACG 985
 DB 601 AAGGGGGCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 986 ATACCGCTGTATGTCTCTGCTGAGACACCTGGCTGACAGACGACGACGACGACG 1033
 DB 661 AAACCGCTGTATGTCTCTGCTGAGACACGACGACGACGACGACGACGACGACG 708
 RESULT 14
 LOCUS B1837271 794 bp mRNA linear EST 04-OCT-2001
 DEFINITION M03090563F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229519 5',
 mRNA sequence.
 ACCESSION B1837271
 VERSION B1837271.1 GI:15948809
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 NIH-MGC http://mgi.nci.nih.gov/
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM1577 row: n column: 16
 High quality sequence stop: 781.
 Location/Qualifiers
 1..794
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5229519"
 /clone_lib="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."
 BASE COUNT 184 a 240 c 195 g 175 t
 ORIGIN
 Query Match 43.3% Score 582.4; DB 13; Length 794;
 Best Local Similarity 98.3%; Pred. No. 1.8e-150;
 Matches 641; Conservative 0; Mismatches 6; Indels 5; Gaps 5;
 QY 1 ATGCCAGGAATAAAAAGGATACCTACTGTTACATTTGCTCTCTGTTCCAAAGCCCT 60
 DB 142 ATGCCAGGAATAAAAAGGATACCTACTGTTACATTTGCTCTCTGTTCCAAAGCCCT 201
 QY 61 GGAATATCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 120
 DB 202 GGAATATCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 261
 QY 121 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

DB 262 TTGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 321
 QY 181 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 DB 322 AACCAAAATGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATG 381
 QY 241 TCGAACCCCTACTGACGACCCCTACTGACGACCCCTACTGACGACCCCTACTGACGAC 300
 DB 382 TCGAACCCCTACTGACGACCCCTACTGACGACCCCTACTGACGACCCCTACTGACGAC 441
 QY 301 GGTCCAAATAT-CCACAGATCTCCAGAGCCTCTTATATGCCCTTTGGATACAGATGA 359
 DB 442 GCTCCAAATATCCACAGATCTCCAGAGCCTCTTATATGCCCTTTGGATACAGATGA 501
 QY 360 TGAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 418
 DB 502 TGAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561
 QY 419 CCACCCAGATCTGATCAATTAATGAAAGCGGGTACACCTGCTCTGACCGACGATATT 478
 DB 562 CCACCCAGATCTGATCAATTAATGAAAGCGGGTACACCTGCTCTGACCGACGATATT 621
 QY 479 GGCCTCTGAGAGGCGCGTCTTACATGATGATGATGATGATGATGATGATGATGATGATG 537
 DB 622 GGCCTCTGAGAGGCGCGTCTTACATGATGATGATGATGATGATGATGATGATGATGATG 681
 QY 538 -CTGTGTGGAATGTT-CCTGATCTTATCTTGTATGATGATGATGATGATGATGATGATG 595
 DB 682 ACTGTGTGGAATGTT-CCTGATCTTATCTTGTATGATGATGATGATGATGATGATGATG 741
 QY 596 ATGAGATGGAAGGCTTGTCCCAAGATGTAACGAGTGTGCCACGAGAACCC 647
 DB 742 ATGAGATGGAAGGCTTGTCCCAAGATGTAACGAGTGTGCCACGAGAACCC 793

RESULT 15
 LOCUS B0716088 933 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT_8099803 Lupskl_sympathetic_trunk Homo sapiens cDNA clone
 IMAGE:6190063 5', mRNA sequence.
 ACCESSION B0716088
 VERSION B0716088.1 GI:21854985
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 933)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM13588 row: e column: 08
 High quality sequence stop: 501.
 Location/Qualifiers
 1..933
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6190063"
 /clone_lib="Lupskl_sympathetic_trunk"
 /sex="male"
 /tissue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:

NotI; Site 2; SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCCCTT-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

BASE COUNT 181 a 266 c 206 g 182 t
ORIGIN

Query Match 46.1%; Score 619; DB 14; Length 835;
Best Local Similarity 98.8%; Pred. No. 1.3e-160;
Matches 645; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

1 ATGCCAGAAATAAAGATACCTGTTACCATCTGCTCTCTCTCCAAAGCCCT 60
|||||
171 ATGCCAGAAATAAAGATACCTGTTACCATCTGCTCTCTCTCCAAAGCCCT 230
61 GGAATGCACAGGACAGTGCAGCAATGGCTTGACCTGCAGTCCAGTACAGTGT 120
|||||
231 GGGATGCACAGGACAGTGCAGCAATGGCTTGACCTGCAGTCCAGTACAGTGT 290
121 TTGATATGATGATATCCGACCATCCCGAGGCTTCCGAGAGACATGATGTGT 180
|||||
291 TTGATATGATGATATCCGACCATCCCGAGGCTTCCGAGAGACATGATGTGT 350
181 AACCAAAATGGGGGATTTATGATCCCGGACCAACCTGTGTATCCAGAGCCCTAC 240
|||||
351 AACCAAAATGGGGGATTTATGATCCCGGACCAACCTGTGTATCCAGAGCCCTAC 410
241 TCGAACCCCTACTGACACCCCTACTGAGTCCGATACCCAGAGTGCCTCCACTCTCA 300
|||||
411 TCGAACCCCTACTGACACCCCTACTGAGTCCGATACCCAGAGTGCCTCCACTCTCA 470
301 GCTCCAAATATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACCATGAT 360
|||||
471 GCTCCAAATATCCACGATCTCCAGGCTCTTATATGCGCTTGGATACCATGAT 530
361 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAACAGATCCACCATGATGCAACCC 420
|||||
531 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAACAGATCCACCATGATGCAACCC 590
421 ACCAGATCTGCATCACTGATGAGAGGCGGTACACCTGCTCAGCAGAGGATTTGG 480
|||||
591 ACCAGATCTGCATCACTGATGAGAGGCGGTACACCTGCTCAGCAGAGGATTTGG 650
481 CTTCTGGAAAGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATG 540
|||||
651 CTTCTGGAAAGGCGGCTTACGATGATGATGATGATGATGATGATGATGATGATG 710
541 TGTGGAAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 599
|||||
711 TGTGGAAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 770
600 GGATGGAAAGGCTTGGCAAGATGTGAACGAGTG--TGCACCGGAAACCCCTG 650
|||||
771 GGATGGAAAGGCTTGGCAAGATGTGAACGAGTG--TGCACCGGAAACCCCTG 823

RESULT 13
BI251219 717 bp mRNA linear EST 17-JUL-2001
LOCUS 602994427F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5150191 5',
DEFINITION mRNA sequence.
ACCESSION BI251219
VERSION BI251219.1 GI:14800410
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 717)
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
Tissue Procurement: Lochar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: ILAM1371 row: e column: 08
High quality sequence stop: 713.

FEATURES
source location/Qualifiers
1..717
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5150191"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lochar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 172 a 186 c 182 g 177 t
ORIGIN

Query Match 44.1%; Score 592.8; DB 13; Length 717;
Best Local Similarity 89.8%; Pred. No. 2.2e-153;
Matches 636; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

326 GGCCCTTATATGCGGCTTTGGATACCAAGATGATGAAGAAACCAACATGTGTGATGTG 385
|||||
1 GGCCCTTATATGCGGCTTTGGATACCAAGATGATGAAGAAACCAACATGTGTGATGTG 60
386 ACAGATGTGCAACAGATTTCCACCAAGTGCACACCCACCAATCTGATCACTGAAG 445
|||||
61 ACAGATGTGCAACAGATTTCCACCAAGTGCACACCCACCAATCTGATCACTGAAG 120
446 GCGGATACCTGCTCTGACACGACGATATGCTTCTGGAAGCCAGTGTGATGACA 505
|||||
121 GAGGTACACCTGCTCTGACACGACGATGCTTCTGGAAGCCAGTGTGATGACA 180
506 TTGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 565
|||||
181 TTGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
566 CTTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTCCAAAGATGTA 625
|||||
241 CTTGTACATGCAACCTGTGTTTACCTCAATGAGATGGAAGTCTTCCAAAGATGTA 300
626 ACAGATGTGCAACAGATTTCCACCAAGTGCACACCCACCAATCTGATCACTGA 685
|||||
301 ACAGATGTGCAACAGATTTCCACCAAGTGCACACCCACCAATCTGATCACTGA 360
686 TCTGCGCTGTGACCCAGATATGATGATGATGATGATGATGATGATGATGATGATG 745
|||||
361 TCTGCGCTGTGACCCAGATATGATGATGATGATGATGATGATGATGATGATGATG 420
746 ACAGATGACACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 805
|||||
421 ACAGATGACACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
806 ACTTGTGCTCTGCTCCAGAGCTACATCCGCTGATGATGATGATGATGATGATGATGAT 865
|||||
481 ACTTGTGCTCTGCTCCAGAGCTACATCCGCTGATGATGATGATGATGATGATGATGAT 540

Db 836 AAA-CTGGCTCAACACCTTACGGATCTTTTCACCTGCGCGTGTGACC 879

RESULT 11
BOJ30927 870 bp mRNA linear EST 20-AUG-2002Z
LOCUS BOJ30927
DEFINITION AGNCNCOURT_8955548 NCL_CGAP_Mam2 Mus musculus cDNA clone
IMAGE:6435732.5, mRNA sequence.
ACCESSION BOJ30927
VERSION BOJ30927.1 GI:22345958
SOURCE EST.
ORGANISM house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 870)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov.
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLMJ3960 Row: d Column: 05
High quality sequence stop: 651.
Location/Qualifiers
1. 870
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6439732"
/clone_id="NCL_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI. Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 200 a 248 c 208 g 212 t 2 others

ORIGIN

Query Match 46.3%; Score 622.6; DB 14; Length 870;
Best Local Similarity 87.7%; Pred. No. 1,3e-161;
Matches 713; Conservative 0; Mismatches 96; Indels 4; Gaps 3;

1 ATGCCAGGAATAAAGAGATACACAGTGTACCATTCCTGGCTCTCTGTTCCAAAGCCCT 60
|||||
48 ATGCCAGGAATTAAGAGATACACAGTGTACCATTCCTGGCTCTCTGTTCCAAAGCCCT 107
|||||
61 GGGAAATGCACAGCAGACAGTGCAGAAATGGCTTTTACCTGGATCGCCAGTCAAGACAGTGT 120
|||||
108 GGGAAATGCACAGCAGACAGTGCAGAAATGGCTTTTACCTGGATCGCCAGTCAAGACAGTGT 167
|||||
121 TTAGATATTGATGAATATGCGGAATCCATCCCGAGGCTGCGGAGGAGACATGATGTGTGT 180
|||||
168 CTAGATATTGATGAATATGCGGAATCCATCCCGAGGCTGCGGAGGAGACATGATGTGTGT 227
|||||
181 AACCAAAATGGCGGGGATTTTATGATATCCCTCCCGGACAAACCCGTGTATCGAGGGCCCTAC 240
|||||
228 AACCAAAATGGCGGGGATTTTATGATATCCCTCCCGGACAAACCCGTGTATCGAGGGCCCTAC 287
|||||
241 TCGAACCCCTACTCGAACCCCTACTCGAGGCGCTATCCACAGCAGTGTGCCCCACACTCTCA 300
|||||
288 TCAATATCCCTACTCTACATCTCTACTCTAGGCGCCATATCCACAGGCGGCCACCACTAGTACA 347
|||||
301 GCTGCAAACTATCCACAGATCTCCAGGCGCTCTTATATGCGGCTTTGGATACCAATGAT 360
|||||

Db 348 GCTTCCACATACCCACGATTTTCAAGCCCTCTGTCTGCGCGCTTTGGGTATCAAGATGAT 407

QY 361 GAAAGCAACCAATGTGTGTGATGTGAGAGCTGTGCAACAGATTTCCACAGTGCACACCC 420

Db 408 GAAGGCAACCAAGTGTGTGATGTGAGAGCTGTGCAACAGATTTCCACAGTGCACACCC 467

QY 421 ACCCAGTCTGATTCGAATACTGAAAGCGGGTAAACCTGCTCTGACACCGAGATATTGG 480

Db 468 ACCCAGTCTGATTCGAATACTGAAAGCGGGTAAACCTGCTCTGACACCGAGATATTGG 527

QY 481 CTTTGTGAAAGCCAGTGTCTTGAACATTTGATGATGTGCTATGTTACTGCTCCAGACGCTC 540

Db 528 CTTTGTGAAAGCGGAGTGTCTTGAACATTTGATGATGTGCTATGTTACTGCTCCAGACGCTC 587

QY 541 TGTGCGCAATGTTCTCTGGATCTCTTCTTGTACATGCAACCCGTGGTTTAAACCTCAATGAG 600

Db 588 TGTGCGCAATGTTCTCTGGATCTCTTCTTGTACATGCAACCCGTGGTTTAAACCTCAATGAG 647

QY 601 GATGGAAGTCTTTGCCAAGATGGAACGAGTGTGCCACGAGAAACCCGTGCGTGCACAC 660

Db 648 GATGGAAGTCTTTGCCAAGATGGAACGAGTGTGCCACGAGAAACCCGTGCGTGCACAC 707

QY 661 TGGCTCAACACTAGGCGCTCTTTCATCTGCGCGCTGTGACCCAGATATG-AACTTGAGA 719

Db 708 TGTGTCAACACTATGCGCTCTTTCATCTGCGCGCTGTGACCCAGATATG-AACTTGAGA 767

QY 720 AGATGGCGTTTCA-TTGCAGTGATATGAGAGAGTGCAGC--TTCTCGAGTTCCTCTGCGCA 776

Db 768 AGATGGCGATTCACNTGTCAGATGATATGAGAGAGTGCAGCCTTCTCCGAGATTTCTCTGTCA 827

QY 777 ACATGAGTGTGGAACCAAGCCCGGCACATACTT 809

Db 828 ACACAGATGTGTGGAACCAAGCCCGGCACATACTT 860

```

RESULT 12
LOCUS      B0718885
DEFINITION B0718885                               835 bp  mRNA          linear  EST 16-JUL-2001
AGENCOCURT_8241132 lupskl_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6187063 5', mRNA sequence.
ACCESSION  B0718885
VERSION    B0718885.1
KEYWORDS   GI:21857782
SOURCE     EST.
           human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 835)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs@mail.nih.gov
           Tissue Procurement: Dr. James R. Lupski
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LNL at:
           http://image.lnl.gov
           plate: LLM13580 row: h column: 08
           High quality sequence stop: 668.
           Location/Qualifiers
             1..835
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone IMAGE:6187063"
               /clone_id="lupskl_sympathetic_trunk"
               /sex="male"
               /tissue_type="sympathetic trunk"
               /dev_stage="adult, 16 yr"
               /lab_host="DH10b"
               /note="Vector: pCMV-Sport6 (Life Technologies); site 1:

```


VERSION	BID53348.1	GI:15244004
KEYWORDS	EST.	house mouse.
SOURCE	Mus musculus.	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	NIH-MGC http://mgc.ncl.nih.gov/.	
AUTHORS	1 (bases 1 to 796)	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bsf-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLML1637 row: f column: 18 High quality sequence stop: 796.	
FEATURES	Location/Qualifiers	
SOURCE	1..796	
	/organism="Mus musculus"	
	/strain="C57BL/6J"	
	/db_xref="taxon:10090"	
	/clone_image:5252369"	
	/clone_lib="NCI_CGAP_Mam5"	
	/tissue_type="tumor, gross tissue"	
	/dev_stage="7 months"	
	/lab_host="DH10B"	
	/note="Organ: mammary; Vector: pCMV-SPORE6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"	
BASE COUNT	195 a 208 c 197 g 196 t	
ORIGIN		
Query Match	47.6%; Score 639.2; DB 13; Length 796;	
Best Local Similarity	89.2%; Pred. No. 3.1e-166;	
Matches	712; Conservative 0; Mismatches 83; Indels 3; Gaps 2	
OY	354 GATGAGTAAGAACAACCAATGTGTGGATGTGGACGAGTGTCACAGATTCCACCAGTg	413
Dd	1 GATGGATTAAAGCAACCAAGTGTGTGGATGTGGACGAGTGTCACAGACTCACACAGTg	60
Y	414 CAACCCCACCCAGATCTCATCAATCAATCAAGAGCGGGTACACCTCTCTGCACCGAGCG	473
Dd	61 CAAACCTTACCCAGATCTGTATCAACATGAAGAGGATPACACTCTCTCTGCACCGAGTg	120
OY	474 ATATTGGCTTTGNAAGGACAGTGTGTAGACATTTGATGAATGTGCTATGTTACTGCCA	533
Dd	121 GTACTGGCTTTCGAAGGACAGTGTGTAGATTTGATGAATGTGCTATGTTACTGCCA	180
OY	534 GCAGCTGTGCGAATGTTCCTGTGATCTTATTCCTGTATCAATCAACCACTGTGTTAACCT	593
Dd	181 GCAGCTGTGCGAATGTTCCTGTGATCTTATTCCTGTATCAACCACTGTGTTAACCT	240
OY	594 CAATGAGATGGAAGGCTTGTGCGAAGATGTGAACAGATGTGTCACCGAGAACCCCTGCT	653
Dd	241 CAAAGCAGATGGAAGGCTTGTGCGAAGATGTGAACAGATGTGTCACCGAGAACCCCTGCT	300
OY	654 GAAACGCTGCTCAACACCTAGAGGCTTTTCATCTGCGCGTGTGACCCAGATATGAAT	713
Dd	301 TCAGACCTGTGTCAACACCTATAGGCTTTTCATCTGCGCGTGTGACCCAGATATGAAT	360
OY	714 TGAGGAAGATGGCTTCAATTCAGTAGATATGAGACGAGTGCAGCTTCTCGATTTCTCTG	773
Dd	361 TGAGGAAGATGGCTTCAATTCAGTAGATATGAGACGAGTGCAGCTTCTCGATTTCTCTG	420
OY	774 CCAACATGATGTGTGAACAGCCCGGACATATCTTGTGCTCTGCCCCCTCAGGCTCAT	833

Db	421	TCAACACGAGGTGTGAACCAAGCCGGGCTCAATTCTTCTGCTGTCGCTCCAGGCTACGT	480		
Qy	834	CCTGCTGGATGACACCGAAGCTGCCAAGACATCAACGAATGTGAGCAGAGAACAAC	893		
Db	481	CCTGTGGATGATATACCGAAGCTGCCAGATATCAATGAATGTGAGCAGACCGAACAAC	540		
Qy	894	GTCGAACCTGACGACGACGTGCTACATTTTACAAAGGGGGCTTCAAATGATCAACCCCAT	953		
Db	541	GTTGACTCTCAATGAGACCTGGCTCAATCTACAAAGGGGGGCTTCAAATGATTTGATCCCAT	600		
Qy	954	CCGCTGTGAGAGGCGCTTATCTGAGGATGAGATTAACCGGCTGTATGTGTCCTGTGAG-A	1012		
Db	601	CAGCTGTGAGAGGCGCTTATCTGCGATTTGGTGAAGAACCGCTGTATGTGTCCTGTGAGCA	660		
Db	661	ACACGACGATGAGAGACCAAGCATTCACATCCTGTATCGGAGACATGSAATGTGTCTAG	720		
Qy	1013	ACCGCTGCGTGCAGACCAAGCCCTTTACATCTTTGTAACCGGGACATGSACTGTGTGAG	1072		
Db	721	GACGCTCCGTTCCGCTGACATCTTCCAAATGCAAGCAGACCCGCTACCGTGGGCGCT	1132		
Qy	1133	ATTACATTTTCCAGATCA	1150		
Db	779	ATTACATTTTCCAGATCA	796		
RESULT 9	BG938621	644 bp	mRNA	linear	EST 20-JUN-2002
LOCUS	BG938621	cn27c11.x1	Normal Human Trabecular Bone Cells	Homo sapiens	cdNA
DEFINITION	cn27c11.x1	Normal Human Trabecular Bone Cells	Homo sapiens	cdNA	sequence.
ACCESSION	BG938621				
VERSION	BG938621.1	GI:14337993			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Beckstron-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.				
TITLE	SGAP: The Skeletal Genome Anatomy Project				
JOURNAL	unpublished (1997)				
COMMENT	Contact: libin jia				
	Medical Genetics Branch				
	National Human Genome Research Institute				
	10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA				
	Tel: 301-402-4877				
	Fax: 301-496-7157				
	Email: libin@helix.nih.gov				
	DNA Sequencing and analyses by National Institutes of Health				
	Intramural Sequencing Center (NISC).				
	Plate: 27 row: G column: 11				
	Seq primer: -21M13 forward primer (ABI).				
FEATURES	Location/Qualifiers				
source	1..644				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="NHBC.cn27c11"				
	/clone_lib="Normal Human Trabecular Bone Cells"				
	/sex="Female"				
	/tissue_type="Bone"				
	/cell_type="Trabecular Bone Cells"				
	/lab_host="SURE"				
	/note="Organ: Hip; Vector: pBluescript; Site:1; ECORI; library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)."				
BASE COUNT	146 a	195 c	154 g	149 t	
ORIGIN					
Query Match	47.3%	Score 636;	DB 13;	Length 644;	

```

Db      185 GAGTCGACGCTTCTGTGAGTCTCTGTGCAACATGATGTGTGAACACGACCGGACATATAC 244
QY      808 TTCTGCTCCTGCCTCCAGGCTACATCTCTGTGATGACACCAAGCTGCCAAGACATC 867
Db      245 TTCTGCTCCTGCCTCCAGGCTACATCTCTGTGATGACACCAAGCTGCCAAGACATC 304
QY      868 AACGAATGTGAGCAGACAGAACCAACACAGCTGCAACCTGACAGACAGCTGTACAAATTTACAA 927
Db      305 AACGAATGTGAGCAGACAGAACCAACACAGCTGCAACCTGACAGACAGCTGTACAAATTTACAA 364
QY      928 GGGGGCTTAAATGATCGACACCCCATCCGCTGTGAGAGAGCTTATCTAGATCATGTAT 987
Db      365 GGGGGCTTAAATGATCGACACCCCATCCGCTGTGAGAGAGCTTATCTAGATCATGTAT 424
QY      988 AACCCGCTATGTCTCTGTGAGAACCTGTGCTGACAGACAGCCCTTTACATCTTG 1047
Db      425 AACCCGCTATGTCTCTGTGAGAACCTGTGCTGACAGACAGCCCTTTACATCTTG 484
QY      1048 TACCGGGACATGAGAGCTGTGTGACAGAGCTCCGCTCCGCTGACATCTTCCAAATGCA 1107
Db      485 TACCGGGACATGAGAGCTGTGTGACAGAGCTCCGCTCCGCTGACATCTTCCAAATGCA 544
QY      1108 GCGAGACACCGGCTGTGAGAGCTGTGTGACAGAGCTCCGCTCCGCTGACATCTTCCAAATGCA 1165
Db      545 GCGAGACACCGGCTGTGAGAGCTGTGTGACAGAGCTCCGCTCCGCTGACATCTTCCAAATGCA 603
QY      1166 GCGAGACATTTTACATGCGGCAACAGGCGCCCATCATGCTGACAGCTGTGTGACAGAGCT 1225
Db      604 GCGAGACATTTTACATGCGGCAACAGGCGCCCATCATGCTGACAGCTGTGTGACAGAGCT 663
QY      1226 CCATCAAA-GGGCCCCGGGAAATCCAGCTGTGAAATGATGATCTGTCAACACTGTC 1284
Db      664 CCATCAAAAGGGGGCGGCTGTGAAATGATGATGATGATGATGATGATGATGATGATGAT 1342
QY      1285 ATCAACTTCAGAG 1297
Db      722 ATCAATTCAGAG 734

```

```

RESULT 7      882 bp      mRNA      linear      EST 04-OCT-2001
LOCUS      BI825121      603038662F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5'
DEFINITION 603038662F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5',
            mRNA sequence.
ACCESSION  BI825121      GI:15936671
VERSION     BI825121.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 882)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cga@bbs-remail.nih.gov
            Tissue procurement: Life Technologies, Inc.
            cDNA library preparation: Life Technologies, Inc.
            DNA sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNML at:
            http://image.llnl.gov
            Plate: LAM11447 row: 1 column: 15
            High quality sequence start: 14
            High quality sequence stop: 844.
            Location/Qualifiers

```

```

FEATURES
    source
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5179550"
        /clone_lib="NIH_MGC_115"

```

```

/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pcmv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC Library."

```

```

BASE COUNT    191 a      277 c      216 g      197 t      1 others
ORIGIN
Query Match    47.7%; Score 641.2; DB 13; Length 882;
Best Local Similarity 98.2%; Pred. No. 9.2e-167;
Matches 690; Conservative 0; Mismatches 9; Indels 4; Gaps 4;
QY      1 ATGCCAGAAATTAAGAGATACATCTGTATTCATCTGTGCTCTGTCTTCCAAAGCCCT 60
Db      179 ATGCCAGAAATTAAGAGATACATCTGTATTCATCTGTGCTCTGTCTTCCAAAGCCCT 238
QY      61 GGGATGACACAGGACAGTGCAGCAATGGCTTTGACCTGTGATCCGACAGTACAGATGT 120
Db      239 GGGATGACACAGGACAGTGCAGCAATGGCTTTGACCTGTGATCCGACAGTACAGATGT 298
QY      121 TTGATATTTGATGATATCCGAAACCATCCGAGGCTCCGAGAGACATGATGTGT 180
Db      299 TTGATATTTGATGATATCCGAAACCATCCGAGGCTCCGAGAGACATGATGTGT 358
QY      181 AACCAAAATGGGGGATTTATGATTTCCCGACAAACCTGTGTATGAGGGCCCTAC 240
Db      359 AACCAAAATGGGGGATTTATGATTTCCCGACAAACCTGTGTATGAGGGCCCTAC 418
QY      241 TCGAACCCTTACCTGAGAGCTTACAGTCCGATCCGACAGAGTCCGACAGTCCGACAGT 300
Db      419 TCGAACCCTTACCTGAGAGCTTACAGTCCGATCCGACAGAGTCCGACAGTCCGACAGT 478
QY      301 GCTCAAACTATGCTCCAGATCTCCAGGCTCTTATATCCGCTTGTGATACAGATGAT 360
Db      479 GCTCAAACTATGCTCCAGATCTCCAGGCTCTTATATCCGCTTGTGATACAGATGAT 538
QY      361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGACACAGATTTCCACAGTGCACACCC 420
Db      539 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGACACAGATTTCCACAGTGCACACCC 598
QY      421 ACCCAATGTGATCATATGAGAGGGGGGATACACCTGCTCTGACAGGAGATATG 480
Db      599 ACCCAATGTGATCATATGAGAGGGGGGATACACCTGCTCTGACAGGAGATATG 658
QY      481 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGCTGTATGTTACTGCGAGAGCTC 540
Db      659 CTTCGGAAGGCGAGTCTTAGACATTTGATGATGCTGTATGTTACTGCGAGAGCTC 718
QY      541 TGTGCGAATTTCTGATCTTATTTGTATGATGACAAACCT-GGTTTACCCTGAAGA 599
Db      719 TGTGCGAATTTCTGATCTTATTTGTATGATGACAAACCTGGGTTTACCCTGAAGA 778
QY      600 GAGTGG-AAGGTCTGCGAAGATGTGAAG-AGTGTGCGAGCAAGACCCCGCTGCA 657
Db      779 GAGTGGCAAGGCTCTTCCCAAGATGTGAAGAGTGTTCACCGAAGAACCTGCGTGC 838
QY      658 ACCTGGGTCAACACCTTACGCTTTTATGATGCGCGCTGTGAC 700
Db      839 AACTGGGTCAACACCTA-GGGTCTTATGATGCGCGCTGTGAC 880

```

```

RESULT 8      796 bp      mRNA      linear      EST 21-AUG-2001
LOCUS      BI453348      60313067F1 NC1_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5252369 5',
DEFINITION 60313067F1 NC1_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5252369 5',
            mRNA sequence.
ACCESSION  BI453348

```

TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

BASE COUNT	222 a	258 c	230 g	181 t
ORIGIN				

Query Match	49.6%	Score 666.8;	DB 12;	Length 891;
Best Local Similarity	95.1%	Pred. No. 7.3e-174;		
Matches 742; Conservative	0;	Mismatches 32;	Indels 6;	Gaps 5

QY	304	CCAAACTATCCACGATCTCCAGGCGCTCTTATATGCCGCTTTTGATACAGATGATGAA	36:
Db	1	CCAAACTATCCACGATCTCCAGGCGCTCTTATATGCCGC-TTGATACAGATGATGAA	59
QY	364	AGCAACCAATGTGTGGATGTGGACGAGTGTGCAACGATTTCCACAGTGCACACCCACC	422
Db	60	AGCAACCAATGTGTGGATGTGGACGAGTGTGCAACGATTTCCACAGTGCACACCCACC	111
QY	424	CAGATCTGCATCATCTGAAAGGCGGGTACACCTGCTCCTGCACGACGATATTGACTT	483
Db	120	CAGATCTGCATCATCTGAAAGGCGGGTACACCTGCTCCTGCACGACGATATTGACTT	176
QY	484	CTGGAAGGCCAGTGGTTGACATGATGAATCCGTATAGTGTACTGACGACGCTGTGT	543
Db	179	CTGGAAGGCCAGTGGTTGACATGATGAATCCGTATAGTGTACTGACGACGCTGTGT	238
QY	544	GCGAATGTTCTGTGATCTTATTTGTACATCAACCTGTGTTTACCCGCAATGAGAT	603
Db	239	GCGAATGTTCTGTGATCTTATTTGTACATCAACCCGTGTTTACCCGCAATGAGAT	298
QY	604	GGAAGGCTTGCCCAAGATGTGACGAGTGTGCCACCGAGAACCCCTGCGTCAAACTGC	663
Db	299	GGAAGGCTTGCCCAAGATGTGACGAGTGTGCCACCGAGAACCCCTGCGTCAAACTGC	358
QY	664	GTCAAACACTACGAGCTCTTTCATCTGCAGCGTGTGACCCGAGATGATCACTTGAGGAAT	723
Db	359	GTCAAACACTACGAGCTCTTTCATCTGCAGCGTGTGACCCGAGATGATCACTTGAGGAAT	418
QY	724	GCGGTTATTTGCAATGATATGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGAG	783
Db	419	GCGGTTATTTGCAATGATATGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGAG	478
QY	784	TGTGTGAACAGCGCCGGCACATACTTGTGCTCCCTCCAGGCTACATCTGTGTGAT	843
Db	479	TGTGTGAACAGCGCCGGCACATACTTGTGCTCCCTCCAGGCTACATCTGTGTGAT	538
QY	844	GACAAACGAACTGCGCAAGACATCAAGCAATGTGACACAGGAAACACAGTGCACACTG	903
Db	539	GACAAACGAACTGCGCAAGACATCAAGCAATGTGACACAGGAAACACAGTGCACACTG	598
QY	904	CAGCAACGCTGCTACATTTTACAAGGGGGCTTCAAA-TGCATCGACCCCATCGCTGTGA	962

Db 599 CAGCAGAGCTGGTACAAATTACAAAGGGGGCTTCAAGATGCAATGCCCATCCGCTGTGA 658

QY 963 GGAGCCTTATCTGAGGATCACTGATACCGCT -GTATGTGTCCTGCTGAGAACCCTTGCT 1021

Db 659 GGAGCCTTATCTGAGGATCACTGATACCGCTGGAATGTGTCCTGCTGAGAAC -TGGT 716

QY 1022 GCAGGACCAAGCCCTTACCATCTTGTACCGGGACATGAGAGCTGGTGCAGAGACGCTCCG 1081

Db 717 GCGAATACAGAGCCCTTACCATCTGTGCTCCGGACATGAGGCCCTGGGGCAGAGACCGGGGCG 776

RESULT 6	BG109391	822 bp	mRNA	linear	EST 30-JAN-2001
LOCUS	BG109391				
DEFINITION	602280448F1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4368227 5',				
	mRNA sequence.				
ACCESSION	BG109391				
VERSION	BG109391.1	GI:12602897			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

Plate: L1AM10021 row: 0 column: 12
High quality sequence stop: 676.

Query Match	49.5%	Score 665;	DB 12;	Length 822;
Best Local Similarity	97.8%;	Pred. No. 2.2e-173;		
Matches 717; Conservative	0;	Mismatches 10;	Indels 6;	Gaps 4;

QY 568 TGAACATGCACACCCCTGGTTTACCCTCAATGAGATGGAAAGGCTTCGCCAAGATGTGCAAC 627

Db 5 TGTACATGCACACCCCTGGTTTACCCTCAATGAGATGGAAAGGCTTCGCCAAGATGTGCAAC 64

QY 628 GAGTGTGCCACGAGAACCCCTGCGTGTGCAAAACCTGGCTCAACACTACGAGCTCTTTCATC 687

Db 65 GAGTGTGCCACGAGAACCCCTGCGTGTGCAAAACCTGGCTCAACACTACGAGCTCTTTCATC 124

QY 688 TGGCGGTGTGAACCCAGATATGAACTTGAAGAAAGATGGCGTTCATTGCACTGATATGAGAC 747

Db 125 TGGCGGTGTGAACCCAGATATGAACTTGAAGAAAGATGGCGTTCATTGCACTGATATGAGAC 184

QY 748 GAGTGCAGCTTCTGTAGTTCTCTGTGCCAACAATGAGTGTGTGNAACAGCCCGACACATAC 807

OY		568	TGTACATGCAACCCTGGTTTAACTTCCTAAGTAGGATGGAAAGCTCTGCACAATGATGAAC	627
Dd		481	TGTACATGCAACCCTGGTTTAACTTCCTAAGTAGGATGGAAAGCTCTGCACAATGATGAAC	540
OY		628	GAGTGTCACACCGAGAACCCTGCGTCACAACTGCGTCACACACTACGGCTCTTTGATC	687
Dd		541	GAGTGTCACACCGAGAACCCTGCGTCACAACTGCGTCACACACTACGGCTCTTTGATC	600
OY		688	TGCCGCTGTGACCCAGCATATGAACCTTGAGGAAATGGCGTTTCATTGCACTGATATGGAC	747
Dd		601	TGCCGCTGTGACCCAGCATATGAACCTTGAGGAAATGGCGTTTCATTGCACTGATATGGAC	660
OY		748	GAGTGCAAGCTTCTGTGAGTTCCTCTGCGCAACATGAGTGTGAACCCAGCCGAGCATAC	807
Dd		661	GAGTGCAAGCTTCTGTGAGTTCCTCTGCGCAACATGAGTGTGAACCCAGCCGAGCATAC	720
OY		808	TTTGTGCTCTGCGCTCCACAGGCTACAT-CCTGCTGGATGATCAACC-GAAGCTGCCAAGA--	863
Dd		721	TTTTGCTCTGCGCTCCACAGGCTACATCCCTGCTGGATGATCAACCAGGCTGCCAAGAAC	780
OY		864	CATCAACGAATGTAGACACAGGAACCAACAGCTGCA--CCTACACAGACCGGTACATTT	922
Dd		781	ATCAACCGAATGTAGACACAGGAACCAACAGCTGCA--CCTACACAGACCGGTACATTT	840
OY		923	T---ACAAGGGGCTTCAATGATCATGCAACCCATCCGCTGTGAGAG	966
Dd		841	TTACAAAGGGGGCTTCAATGATCATGCAACCCATCCGCTGTGAGAG	887
RESULT_4				
BMS545444				
LOCUS				
DEFINITION			954 bp mRNA linear EST 20-FEB-2002	
ACCESSION			AE000001.1	
VERSION			5.1	
KEYWORDS			5' UTR sequence.	
SOURCE			BMS545444	
ORGANISM			BMS545444.1 GI:18777513	
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
SOURCE				
PAGE COUNT				

ORIGIN	Query Match	57.48; Score 771; DB 13; Length 954;
	Best Local Similarity 100.0%; Pred. No. 9.2e-203;	
	Matches 771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	574 TGCAACCTGGTTTATACCTCAATAGAGATGGAAAGTTGTCCCAAGATGTGAAGAGT	633
Db	8 TGCACACCTGGTTTATACCTCAATAGAGATGGAAAGTTGTCCCAAGATGTGAAGAGT	67
OY	634 GCCACCGAGAACCCCTGGCTGCAAAACCTGCTCAACACCTACGGCTCTTTCATCTGC	693
Db	68 GCCACCGAGAACCCCTGGCTGCAAAACCTGCTCAACACCTACGGCTCTTTCATCTGC	127
OY	694 TGTGACCCAGATATGAACCTTGAGAGAGATGGCTTCACTTCAGTATATGACAGATG	753
Db	128 TGTGACCCAGATATGAACCTTGAGAGAGATGGCTTCACTTCAGTATATGACAGATG	187
OY	754 AGCTTCTCTGAGTTCCTCTGCAACATAGTGTGTGAACACGACCGGACATACCTTCTG	813
Db	188 AGCTTCTCTGAGTTCCTCTGCAACATAGTGTGTGAACACGACCGGACATACCTTCTG	247
OY	814 TCTTGCCTCCAGGCTACATCTCTGTGATCAACACGAGACCTGCCAAGATCAAGAA	873
Db	248 TCTTGCCTCCAGGCTACATCTCTGTGATCAACACGAGACCTGCCAAGATCAAGAA	307
OY	874 TGTGAGCAGAGAACACACGCTGCAACCTTGAGAGAGACGCTCAATTTCAAGGGGGC	933
Db	308 TGTGAGCAGAGAACACACGCTGCAACCTTGAGAGAGACGCTCAATTTCAAGGGGGC	367
OY	934 TTCAAAATGATCGAACCCCATCCGCTGTGAGAGACCTTATGTGAGATCAGTATACCG	993
Db	368 TTCAAAATGATCGAACCCCATCCGCTGTGAGAGACCTTATGTGAGATCAGTATACCG	427
OY	994 TGTATGTGCTCGTGAGAACCTTGCGGTGAGAGACGACCCCTTACCATTTGACCGG	1053
Db	428 TGTATGTGCTCGTGAGAACCTTGCGGTGAGAGACGACCCCTTACCATTTGACCGG	487
OY	1054 GACATGAGACGAGTGTCTGAGACGCTCCGCTTCCGCTGACATCTTCCAATGCAAGCCAG	1113
Db	488 GACATGAGACGAGTGTCTGAGACGCTCCGCTTCCGCTGACATCTTCCAATGCAAGCCAG	547
OY	1114 ACCCGCTACCTGGGGCTATTACATTTTCCAGATCAATTTGGGAATGAGGCGAGAGAA	1173
Db	548 ACCCGCTACCTGGGGCTATTACATTTTCCAGATCAATTTGGGAATGAGGCGAGAGAA	607
OY	1174 TTTTACATGCGGCAAGCGGGCCCATCATGCTCCACCCGCGTATGACACGGCCCATCAA	1233
Db	608 TTTTACATGCGGCAAGCGGGCCCATCATGCTCCACCCGCGTATGACACGGCCCATCAA	667
OY	1234 GGGCCCCGGGAAATCCAGCTTGAGCTTGGAAATGATCACTGTCAACACTGTCAATCTT	1293
OY	1294 AGAGGAGAGCTCGTATCCGACTCGGCTGGGATATATGTCGGAGTACCCATTC	1344
Db	668 GGGCCCCGGGAAATCCAGCTTGAGCTTGGAAATGATCACTGTCAACACTGTCAATCTT	727
Db	728 AGAGGAGAGCTCGTATCCGACTCGGCTGGGATATATGTCGGAGTACCCATTC	778
RESULT 5		
BF339448	891 bp	mRNA linear EST 22-NOV-2000
LOCUS	602039102P1	NCI_Gap_Brn64 Homo sapiens cDNA clone IMAGE:4186704
DEFINITION	5', mRNA sequence.	
ACCESSION	BF339448	
VERSION	BF339448.1	GI:11285903
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 891)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 1982.41 Seconds
(Without alignments)
10979.953 Million cell updates/sec

Title: US-09-674-379a-11

Perfect-score: 1344
Sequence: 1 atgccagataaaagat.....atgttcagcagatccattc 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	866	64.4	1014	14	BM921371 AGENCOURT
2	825.4	61.4	996	14	BQ720516 AGENCOURT
3	792.2	58.9	904	14	BQ716569 AGENCOURT
4	771	57.4	954	13	BM545444 AGENCOURT
5	666.8	49.6	891	12	BF339448 AGENCOURT
6	665	49.5	822	12	BG109391 BG109391 602280448

7	641.2	47.7	882	13	BI825121	BI825121 603038662
8	639.2	47.6	796	13	BI453348	BI453348 603173067
9	636	47.3	644	13	BG938621	BG938621 cn27c11.x
10	624.6	46.5	880	13	BI764020	BI764020 603043221
11	622.6	46.3	870	14	BQ930927	BQ930927 AGENCOURT
12	619	46.1	835	14	BQ718885	BQ718885 AGENCOURT
13	592.8	44.1	717	13	BI251219	BI251219 602994427
14	582.4	43.3	794	13	BI837271	BI837271 603090563
15	576.4	42.9	933	14	BQ716088	BQ716088 AGENCOURT
16	573.6	42.7	612	13	BF090699	BF090699 QV2-N7004
17	564.4	42.0	912	13	BI768516	BI768516 603057118
18	555	41.3	721	13	BI523975	BI523975 603052129
19	553.2	41.2	815	13	BG174360	BG174360 602336371
20	540.2	40.2	627	13	BI183862	BI183862 UN2-P-FN-
21	537.8	40.0	745	12	BG431026	BG431026 602498636
22	534.4	39.8	1006	9	AL545549	AL545549 AL545549
23	521.2	38.8	762	13	BI834712	BI834712 603090304
24	510	37.9	531	10	AW957458	AW957458 EST369423
25	486	36.2	889	14	AL041285	AL041285 DKFZP434M
26	480.4	35.7	889	14	BQ897144	BQ897144 AGENCOURT
27	479	35.6	491	14	BM707581	BM707581 UI-E-C11-
28	477.8	35.6	494	14	BQ695207	BQ695207 1001350 H
29	456.2	33.9	685	9	AL543659	AL543659 AL543659
30	452	33.6	629	13	BI489863	BI489863 603031610
31	449.2	33.4	1015	13	BM544363	BM544363 AGENCOURT
32	442	32.9	690	12	BF104873	BF104873 601822595
33	434.2	32.3	506	14	BQ723059	BQ723059 AGENCOURT
34	415.4	30.9	576	10	AW142226	AW142226 EST29462
35	409.8	30.5	770	12	BG542457	BG542457 602569534
36	391.2	29.1	680	10	AW106432	AW106432 um27g01.Y
37	389.6	29.0	1101	12	BG245848	BG245848 602358649
38	387.8	28.9	391	14	BM699401	BM699401 UI-E-DX1-
39	384.2	28.6	462	10	BE656312	BE656312 UI-E-BHO-
40	382.4	28.5	719	12	BG718928	BG718928 602692222
41	375.4	27.9	377	14	N95751	N95751 zb55e06.s1
42	371.2	27.6	1850	11	BC018871	BC018871 Homo sapi
43	370.4	27.6	389	10	BE004365	BE004365 CMO-BN010
44	368.8	27.4	858	9	AL546863	AL546863 AL546863
45	366.6	27.3	430	10	AW744572	AW744572 UR32n12.Y

ALIGNMENTS

RESULT 1
LOCUS BM921371 1014 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
ACCESSION BM921371
VERSION BM921371.1 GI:19371750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1014)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM12787 row: 1 column: 10
High quality sequence stop: 738.
Location/Qualifiers 1..1014

FEATURES

source

```
QY 481 CTTCTGGAAGGCCAGTCTTATGACATTTGATGTAATGTGCTATGTTACTGACAGACTC 540
   |||||||
Db 930 CTTCTGGAAGGCCAGTCTTATGACATTTGATGTAATGTGCTATGTTACTGACAGACTC 989
QY 541 TGTGCGAATGTTCTGTGATCCATTTCTTTGATCATGCAACCCGTGTTTACCCCTCAATGAG 600
   |||||||
Db 990 TGTGCGAATGTTCTGTGATCCATTTCTTTGATCATGCAACCCGTGTTTACCCCTCAATGAG 1049
QY 601 GATGGAAGTCTTGTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAACC 660
   |||||||
Db 1050 GATGGAAGTCTTGTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCAACC 1109
QY 661 TGGCTCAACACTACGGCTCTTTTCATCTGCGCTGTGACCCGAGATATGACTTGGAGAA 720
   |||||||
Db 1110 TGGCTCAACACTACGGCTCTCTCATCTGCGCTGTGACCCGAGATATGACTTGGAGAA 1169
QY 721 GATGCGTTTCATTCAGTATATGAGACAGTGTGACAGCTTCTGTAGTTCTCTGCAACAT 780
   |||||||
Db 1170 GATGCGTTTCATTCAGTATATGAGACAGTGTGACAGCTTCTGTAGTTCTCTGCAACAT 1229
QY 781 GAGTGTGTAACCAAGCCCGGACATCTTCTGCTCTGCTCCCTCCAGGCTACATCCGCTG 840
   |||||||
Db 1230 GAGTGTGTAACCAAGCCCGGACATCTTCTGCTCTGCTCCCTCCAGGCTACATCCGCTG 1289
QY 841 GATGACAACCGAAGCTGCCAAGACATCAACGATGTGAGCACAGGACACACAGTGCAC 900
   |||||||
Db 1290 GATGACAACCGAAGCTGCCAAGACATCAACGATGTGAGCACAGGACACACAGTGCAC 1349
QY 901 CTGCGACAGACGTGCTCAATTTACAAAGGGGCTTCAAAATGCATGACCCCATCCGCTGT 960
   |||||||
Db 1350 CTGCGACAGACGTGCTCAATTTACAAAGGGGCTTCAAAATGCATGACCCCATCCGCTGT 1409
QY 961 GAGGAGCTTATCTGAGGATCAGATGATACCGCTGATGCTGCTGAGAAACCTGGC 1020
   |||||||
Db 1410 GAGGAGCTTATCTGAGGATCAGATGATACCGCTGATGCTGCTGAGAAACCTGGC 1469
QY 1021 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGACATGAGAGTGTGCAGAGACGCTCC 1080
   |||||||
Db 1470 TGCAGAGACCAAGCCCTTACCATCTTGTACCGGACATGAGAGTGTGCAGAGACGCTCC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGCGGCTATTACATT 1140
   |||||||
Db 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCACGACCCGCTACCTGCGGCTATTACATT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
   |||||||
Db 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1649
QY 1201 AGTGCCACCTGTGTATGACACGCCCATCAAAAGGGCCCGGGAATCCAGCTGACTTG 1260
   |||||||
Db 1650 AGTGCCACCTGTGTATGACACGCCCATCAAAAGGGCCCGGGAATCCAGCTGACTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGACAGTCCGTGATCCGACTGCGG 1320
   |||||||
Db 1710 GAAATGATCACTGTCAACACTGTTCATCAACTTCAGAGGACAGTCCGTGATCCGACTGCGG 1769
QY 1321 ATATATGTGTGAGTACCATTC 1344
   |||||||
Db 1770 ATATATGTGTGAGTACCATTC 1793
   |||||||
```

Search completed: July 3, 2003, 17:53:35
Job time : 221.968 secs

```

QY 181 AACCAAAATGGCGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATGAGGGCCCTAC 240
Db 630 AACCAAAATGGCGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATGAGGGCCCTAC 689
QY 241 TCGAATCCCTACGTGAGACCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 300
Db 690 TCGAATCCCTACGTGAGACCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 749
QY 301 GCTCCAAATATCCAGAGATCTCCAGGCGCTTATATGCGGCTTGGATACAGATGAT 360
Db 750 GCTCCAAATATCCAGAGATCTCCAGGCGCTTATATGCGGCTTGGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACCAATGCAACCC 420
Db 810 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACCAATGCAACCC 869
QY 421 ACCAGATCTGCATCAATGAGAGGGGGGTACACCTCTCTCTGACCGAGGATATG 480
Db 870 ACCAGATCTGCATCAATGAGAGGGGGGTACACCTCTCTCTGACCGAGGATATG 929
QY 481 CTCTGTGAGAGGCGAGTCTTACATTTGATGATGTGCTATGTTACTGCCAGCAGCTC 540
Db 930 CTCTGTGAGAGGCGAGTCTTACATTTGATGATGTGCTATGTTACTGCCAGCAGCTC 989
QY 541 TGTGCGAATGTTCTGTGATCTTATCTTGTATGATGCAACCCCTGTTTACCTCAATGAG 600
Db 990 TGTGCGAATGTTCTGTGATCTTATCTTGTATGATGCAACCCCTGTTTACCTCAATGAG 1049
QY 601 GATGAGAGGTCTTGCAGATGTGCAAGAGTGTGCCACCGAGAACCCCTGGGTGCAACC 660
Db 1050 GATGAGAGGTCTTGCAGATGTGCAAGAGTGTGCCACCGAGAACCCCTGGGTGCAACC 1109
QY 661 TGGCTTAACACCTTACGCGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTGAGGAA 720
Db 1110 TGGCTTAACACCTTACGCGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTGAGGAA 1169
QY 721 GATGGGCTTATGAGTGTATGAGAGAGTGTGAGAGCTTCTGTGATGTTCTTCCCAAT 780
Db 1170 GATGGGCTTATGAGTGTATGAGAGAGTGTGAGAGCTTCTGTGATGTTCTTCCCAAT 1229
QY 781 GAGTGTGTGACACCGCCGACATATCTTGTCTGCTGCGCTTCCAGGCTACATCTGCTG 840
Db 1230 GAGTGTGTGACACCGCCGACATATCTTGTCTGCTGCGCTTCCAGGCTACATCTGCTG 1289
QY 841 GATGACAAACCGAGCTGCCAGACATCAACGAATGTGAGACAGAGAACACACGCTGCAAC 900
Db 1290 GATGACAAACCGAGCTGCCAGACATCAACGAATGTGAGACAGAGAACACACGCTGCAAC 1349
QY 901 CTGCAAGACAGGTGTACATTTTACAAAGGGGGCTTCAAAATGATCGAACCCCATCCGCTG 960
Db 1350 CTGCAAGACAGGTGTACATTTTACAAAGGGGGCTTCAAAATGATCGAACCCCATCCGCTG 1409
QY 961 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAC 1020
Db 1410 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAC 1469
QY 1021 TGCAGAGACCAAGCCCTTACCAATCTTGTACCGGGAATGAGAGCTGTGAGACGCTCC 1080
Db 1470 TGCAGAGACCAAGCCCTTACCAATCTTGTACCGGGAATGAGAGCTGTGAGACGCTCC 1529
QY 1081 GTTCCCGTGTGATCTTCCAAATGCAAGCAGACCGGCTTACCTGGGGCTTATTAATT 1140
Db 1530 GTTCCCGTGTGATCTTCCAAATGCAAGCAGACCGGCTTACCTGGGGCTTATTAATT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGGCCCATC 1200
Db 1590 TTCCAGATCAAAATCTGGGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGGCCCATC 1649
QY 1201 AGTGCACCCCTGTGTATGACAGCCCATCAAAAGGGGCCCATCAAGCTGTGAGACTTG 1260
Db 1650 AGTGCACCCCTGTGTATGACAGCCCATCAAAAGGGGCCCATCAAGCTGTGAGACTTG 1709
QY 1261 GAAATGATCACTGTGATCAACTGTCAAGGAGGAGGCGTCCGTGATCCGACTGCGG 1320

```

```

Db 1710 GAAATGATCACTGTGATCAACTGTCAACTTCAAGAGCAAGCTCCGTGATCCGACTGCGG 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
Db 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 15
US-10-180-552-407
; Sequence 407, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-552-407

Query Match          99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGAAATATAAAGATACCTGATTTACATTTGCTGCTCTGCTTCCAGGCCCT 60
Db 450 ATGCCAGAAATATAAAGATACCTGATTTACATTTGCTGCTCTGCTTCCAGGCCCT 509
QY 61 GGGAAATGCACAGGACAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 120
Db 510 GGGAAATGCACAGGACAGTGTGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 569
QY 121 TTAGATATTTGATGAATGCGGAACCATCCCGAGGCGCTGCCGAGAGAGATGATGTGTT 180
Db 570 TTAGATATTTGATGAATGCGGAACCATCCCGAGGCGCTGCCGAGAGAGATGATGTGTT 629
QY 181 AACCAAAATGGGGGGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATGAGGGCCCTAC 240
Db 630 AACCAAAATGGGGGGGTATTATGCAATTCGCCGAGCAAAACCTGTGTATGAGGGCCCTAC 689
QY 241 TCGAATCCCTACGTGAGACCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 300
Db 690 TCGAATCCCTACGTGAGACCCCTACGAGTCCGTACCCAGAGCTGCCCCACACTCTCA 749
QY 301 GCTCCAAATATCCAGAGATCTCCAGGCGCTTATATGCGGCTTGGATACAGATGAT 360
Db 750 GCTCCAAATATCCAGAGATCTCCAGGCGCTTATATGCGGCTTGGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACCAATGCAACCC 420
Db 810 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACCAATGCAACCC 869
QY 421 ACCAGATCTGCATCAATGAGAGGGGGGTACACCTCTCTCTGACCGAGGATATG 480
Db 870 ACCAGATCTGCATCAATGAGAGGGGGGTACACCTCTCTCTGACCGAGGATATG 929

```

TYPE: DNA
ORGANISM: Homo Sapien
us-10-176-757-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCCAGAAATAAAGAGATCTACCTGTTACCATCTGCTCTCTCTCCAAAGCCCT 60
DB ATGCCAGAAATAAAGAGATCTACCTGTTACCATCTGCTCTCTCTCTCCAAAGCCCT 509
QY 61 GGAATGACAGGACAGTGCAGCAATGCTTACCTGATGCCAGTGCAGAGAGTGT 120
DB GGAATGACAGGACAGTGCAGCAATGCTTACCTGATGCCAGTGCAGAGAGTGT 569
QY 121 TTGATATTGATGATGCGCAACCATCCCGAGGCTGCGCAGAGAGACATGATGT 180
DB TTGATATTGATGATGCGCAACCATCCCGAGGCTGCGCAGAGAGACATGATGT 629
QY 181 AACCAAAATGGGGGTATTATGATCTCCCGGCAAAACCTGTGTATCGAGGGCCCTAC 240
DB AACCAAAATGGGGGTATTATGATCTCCCGGCAAAACCTGTGTATCGAGGGCCCTAC 689
QY 241 TCGAACCCCTACTCGACCCCTACTAGTCCGTAACCCAGAGCTGCCCACTCTCA 300
DB TCGAACCCCTACTCGACCCCTACTAGTCCGTAACCCAGAGCTGCCCACTCTCA 749
QY 301 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGGCTTGTGATPACAGATGAT 360
DB GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGGCTTGTGATPACAGATGAT 809
QY 361 GAAAGAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACCACTGCAACCC 420
DB GAAAGAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACCACTGCAACCC 869
QY 421 ACCCAGATCTGCATCAATCTGAAGGGGTACACCTGCTCTCTCCAGCAGAGATTTGG 480
DB ACCCAGATCTGCATCAATCTGAAGGGGTACACCTGCTCTCTCCAGCAGAGATTTGG 929
QY 481 CTTCTGGAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB CTTCTGGAGGCGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 989
QY 541 TGTGGAATGTCTCTGATGATCTTCTTGTATGATGCAACCTGTGTTTACCTCAATGAG 600
DB TGTGGAATGTCTCTGATGATCTTCTTGTATGATGCAACCTGTGTTTACCTCAATGAG 1049
QY 601 GATGGAAGGTCTTCCCAAGATGTGAAGAGTGTGCCAAGCAACCCCTGCTGCAAC 660
DB GATGGAAGGTCTTCCCAAGATGTGAAGAGTGTGCCAAGCAACCCCTGCTGCAAC 1109
QY 661 TCCGTCAACACCTGAGGCTCTTCTTATGCTGCGCTGTGACCAAGATATGAACTTGAGGA 720
DB TCCGTCAACACCTGAGGCTCTTCTTATGCTGCGCTGTGACCAAGATATGAACTTGAGGA 1169
QY 721 GATGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB GATGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 781 GAGTGTGTAACCGCCGCGACATATCTTGTGCTGCGCTGCAAGGCTACATCTGTG 840
DB GAGTGTGTAACCGCCGCGACATATCTTGTGCTGCGCTGCAAGGCTACATCTGTG 1289
QY 841 GATGACAAACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATG 900
DB GATGACAAACCGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATG 1349
QY 901 CTGACAGCAGAGCTGTACAAATTTACAAGGGGGCTTCAATGATGATGATGATGATG 960
DB CTGACAGCAGAGCTGTACAAATTTACAAGGGGGCTTCAATGATGATGATGATGATG 1409
QY 961 GAGGAGCCTTATCTAGATCACTGATTAACCGCTGATGTGTCTGCTGAGAAACCTGTGC 1020
```

```
DB 1410 GAGGAGCCTTATCTAGATCACTGATTAACCGCTGATGTGTCTGCTGAGAAACCTGTGC 1469
QY 1021 TGCAGAGACCAAGCCCTTTACCATCTGTTACCGGGACATGAGAGTGTGTACAGACCTGC 1080
DB 1470 TGCAGAGACCAAGCCCTTTACCATCTGTTACCGGGACATGAGAGTGTGTGTACAGACCTGC 1529
QY 1081 GTTCCCGCTGATCTTCCAAATGCAAGCAGACCGGCTACCTGCGGGCTTATCAAT 1140
DB 1530 GTTCCCGCTGATCTTCCAAATGCAAGCAGACCGGCTACCTGCGGGCTTATCAAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATCGGCAAAAGGGCCCCATC 1200
DB 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATCGGCAAAAGGGCCCCATC 1649
QY 1201 AGTCCACCCCTGTGTATGATGACAGCCCATCAATCAAGGGCCCCGGGAATTCAGCTGTG 1260
DB 1650 AGTCCACCCCTGTGTATGATGACAGCCCATCAATCAAGGGCCCCGGGAATTCAGCTGTG 1709
QY 1261 GAAATGATCACTGTCAACACGTGATCACTTCAAGAGGAGCTGCGTGTATGCGACTGCGG 1320
DB 1710 GAAATGATCACTGTCAACACGTGATCACTTCAAGAGGAGCTGCGTGTATGCGACTGCGG 1769
QY 1321 ATATATGTGTGCGAGTACCCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCCATTC 1793
```

RESULT 14

US-10-176-913-407
Sequence 407, Application US/10176913
Publication No. US20030022298A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OR INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C66
CURRENT APPLICATION NUMBER: US/10/176, 913
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See file Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-913-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGCCAGAAATAAAGAGATCTACCTGTTACCATCTGCTCTCTCTCCAAAGCCCT 60
DB ATGCCAGAAATAAAGAGATCTACCTGTTACCATCTGCTCTCTCTCTCCAAAGCCCT 509
QY 61 GGAATGACAGGACAGTGCAGCAATGCTTACCTGATGCCAGTGCAGAGAGTGT 120
DB GGAATGACAGGACAGTGCAGCAATGCTTACCTGATGCCAGTGCAGAGAGTGT 569
QY 121 TTGATATTGATGATGCGCAACCATCCCGAGGCTGCGCAGAGAGACATGATGTGT 180
DB TTGATATTGATGATGCGCAACCATCCCGAGGCTGCGCAGAGAGACATGATGTGT 629
```


APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C86
 CURRENT APPLICATION NUMBER: US/10/176,482
 CURRENT FILING DATE: 2002-06-20
 PRIOR APPLICATION removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 407
 LENGTH: 2609
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-176-482-407

Query Match 99.9%; Score 1342.4; DB 9; Length 2609;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGCCGGAATAAAGGATACATCTGTTACCATTTCTGGCTCTCTGCTTCCAGCCCT 60
 450 ATGCCGGAATAAAGGATACATCTGTTACCATTTCTGGCTCTCTGCTTCCAGCCCT 509
 61 GGGAAATGACAGGACAGAGTGCAGCAATGGCTTTGACCTGATGCGCAATGAGACAGT 120
 510 GGGAAATGACAGGACAGAGTGCAGCAATGGCTTTGACCTGATGCGCAATGAGACAGT 569
 121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 570 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
 181 AACCAAAATGCGGGGATTTATGATTTCCCGGCAAAACCTGTGTATGAGAGGCGCTAC 240
 630 AACCAAAATGCGGGGATTTATGATTTCCCGGCAAAACCTGTGTATGAGAGGCGCTAC 689
 241 TCGAACCCCTACTGAGACCCCTACTAGTCCGATCCAGACAGAGTGCAGGACCTCTCA 300
 690 TCGAACCCCTACTGAGACCCCTACTAGTCCGATCCAGACAGAGTGCAGGACCTCTCA 749
 301 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTGTGATACCATGATGAT 360
 750 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTGTGATACCATGATGAT 809
 361 GAAAGAACCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 810 GAAAGAACCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
 421 ACCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 870 ACCAGATCTGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
 481 CTTTCTGGAAGGCCAGTGTAGACATGATGATGATGATGATGATGATGATGATGATGAT 540
 930 CTTTCTGGAAGGCCAGTGTAGACATGATGATGATGATGATGATGATGATGATGATGAT 989
 541 TGTGGGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 990 TGTGGGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
 601 GATGGAAGGCTTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 1050 GATGGAAGGCTTTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
 661 TGGGTCAACACCTACGCGCTTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAGGAA 720

1110 TGGGTCAACACCTACGCGCTTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAGGAA 1169
 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 1170 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
 781 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 1230 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 841 GATGACAAACGAGCTCCCAAGATCAATGATGATGATGATGATGATGATGATGATGATGAT 900
 1290 GATGACAAACGAGCTCCCAAGATCAATGATGATGATGATGATGATGATGATGATGATGAT 1349
 901 CTGACAGACAGCTGCTCAATTTCAAGAGGCGCTTCAATGATGATGATGATGATGATGAT 960
 1350 CTGACAGACAGCTGCTCAATTTCAAGAGGCGCTTCAATGATGATGATGATGATGATGAT 1409
 961 GAGGAGCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 1410 GAGGAGCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
 1021 TGCAGAGACAGCCCTTTACCATCTGTACCGGGACATGAGCTGTGTGAGAGAGCTTCC 1080
 1470 TGCAGAGACAGCCCTTTACCATCTGTGTACCGGGACATGAGCTGTGTGAGAGAGCTTCC 1529
 1081 GTTCCGCTGATCATCTTCCAAATGCAAGCCAGACCCCTTACCTTGGGGCTTATACAT 1140
 1530 GTTCCGCTGATCATCTTCCAAATGCAAGCCAGACCCCTTACCTTGGGGCTTATACAT 1589
 1141 TTCCAGATCAATCTGGGAATGAGGCGAGAGAAATTTACATGCGGCAAAAGGCGCCATC 1200
 1590 TTCCAGATCAATCTGGGAATGAGGCGAGAGAAATTTACATGCGGCAAAAGGCGCCATC 1649
 1201 AGTGCCACCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 1650 AGTGCCACCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
 1261 GAATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1710 GAATATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
 1321 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
 1770 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793

RESULT 13
 US-10-176-757-407
 Sequence 407, Application US/10176757
 Publication No. US2003002297A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C86
 CURRENT APPLICATION NUMBER: US/10/176,757
 CURRENT FILING DATE: 2002-06-20
 PRIOR APPLICATION removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 407
 LENGTH: 2609

Db 1290 GATGACACCGAAGCTGCCAAGACATCAAGAAATGTGAGCAGAGAACCAACGACGTCAAC 1349
QY CTGACACAGACGCTACATTTTACAAGGGGCTTCAATGATCGAGCCCATCCGCTGT 960
Db 1350 CTGACACAGACGCTACATTTTACAAGGGGCTTCAATGATCGAGCCCATCCGCTGT 1409
QY GAGGAGCCCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1020
Db 1410 GAGGAGCCCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1469
QY 1021 TGCAGAGACCGACCCCTTTTACCATCTTGTACCGGAGATGAGCTGTGTGACGAGCTCC 1080
Db 1470 TGCAGAGACCGACCCCTTTTACCATCTTGTACCGGAGATGAGCTGTGTGACGAGCTCC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCGACCCGCTACCTGGGGGCTTATACAT 1140
Db 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCGACCCGCTACCTGGGGGCTTATACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAAGGGGCCCATC 1200
Db 1590 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAAGGGGCCCATC 1649
QY 1201 AGTGCACCCCTGTGTATGACACGCCCATCAAGGGGCCCGGAAATCCAGCTGACTTG 1260
Db 1650 AGTGCACCCCTGTGTATGACACGCCCATCAAGGGGCCCGGAAATCCAGCTGACTTG 1709
QY 1261 GAAATGATCAGTGTCAACACTGTATGATGAGGAGCTCCGCTGATCGAGCTGGGG 1320
Db 1710 GAAATGATCAGTGTCAACACTGTATGATGAGGAGCTCCGCTGATCGAGCTGGGG 1769
QY 1321 ATATATGTGTGCGAGTACCCATTC 1344
Db 1770 ATATATGTGTGCGAGTACCCATTC 1793

RESULT 10

US-10-175-738-407
Sequence 407, Application US/10175738
Publication No. US2003022294A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jilan
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-19
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-738-407

Query Match 99.98; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;

Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGATATAAAGGATGATCTACATTTGCTGCTCTGCTTCCAGAGCCCT 60
Db 450 ATGCCAGATATAAAGGATGATCTACATTTGCTGCTCTGCTTCCAGAGCCCT 509
QY 61 GGGAAATGACAGGACAGTGCAGGAATGCTTGTACCTGATCGCAGTCAAGACATGT 120

Db 510 GGGAAATGACAGGACAGTGCAGGAATGCTTGTACCTGATCGCAGTCAAGACATGT 569
QY 121 TTGATATTTATGATGATCCGACACCATCCCGAGGCTCCGAGAGACATGATGTGT 180
Db 570 TTGATATTTATGATGATCCGACACCATCCCGAGGCTCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGGGGATTTTATGATCCCGGAGCAAAACCTGTGTATGAGGGGCTTC 240
Db 630 AACCAAAATGGGGGATTTTATGATCCCGGAGCAAAACCTGTGTATGAGGGGCTTC 689
QY 241 TCGAACCCCTACTGACCCCTTACTAGTCCGTAACCGAGAGTGTGCCACACTCTCA 300
Db 690 TCGAACCCCTACTGACCCCTTACTAGTCCGTAACCGAGAGTGTGCCACACTCTCA 749
QY 301 GCTCCAAATATCCACAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 360
Db 750 GCTCCAAATATCCACAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTCACACAGTCAACCC 420
Db 810 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTCACACAGTCAACCC 869
QY 421 ACCAGATCTGATCAATGATGAAGGGGATACACCTGCTCTGACCGAGGATATTGG 480
Db 870 ACCAGATCTGATCAATGATGAAGGGGATACACCTGCTCTGACCGAGGATATTGG 929
QY 481 CTTCTGGAAGGCGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 930 CTTCTGGAAGGCGAGCTCTTATGATGATGATGATGATGATGATGATGATGATGAT 989
QY 541 TGTGCAATGTTCTGTGATCTTATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 990 TGTGCAATGTTCTGTGATCTTATGATGATGATGATGATGATGATGATGATGATGAT 1049
QY 601 GATGAAAGTCTTGCACAGATGTGACAGAGTGTGACACCGAGAACCCCTGCTGCAACC 660
Db 1050 GATGAAAGTCTTGCACAGATGTGACAGAGTGTGACACCGAGAACCCCTGCTGCAACC 1109
QY 661 TGGGTCAACACCTACAGGCTCTTATGATGATGATGATGATGATGATGATGATGAT 720
Db 1110 TGGGTCAACACCTACAGGCTCTTATGATGATGATGATGATGATGATGATGATGAT 1169
QY 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 1170 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
QY 781 GAGTGTGTGAACAGCCCGGACATCTTGTCTGCTGCTCCAGGCTACATCTGCTG 840
Db 1230 GAGTGTGTGAACAGCCCGGACATCTTGTCTGCTGCTCCAGGCTACATCTGCTG 1289
QY 841 GATGACACCGGAAGCTCCCAAGACATCAACGAATGTGAGCACAGGAACACACGTCGAC 900
Db 1290 GATGACACCGGAAGCTCCCAAGACATCAACGAATGTGAGCACAGGAACACACGTCGAC 1349
QY 901 CTGACAGAGAGTGTCTCAATTTTACAAGGGGCTTCAATGATGATGATGATGATGAT 960
Db 1350 CTGACAGAGAGTGTCTCAATTTTACAAGGGGCTTCAATGATGATGATGATGATGAT 1409
QY 961 GAGGAGCCCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1020
Db 1410 GAGGAGCCCTTATCTGAGGATCAGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1469
QY 1021 TGCAGAGACCGACCCCTTTTACCATCTTGTACCGGAGATGAGCTGTGTGACGAGCTCC 1080
Db 1470 TGCAGAGACCGACCCCTTTTACCATCTTGTACCGGAGATGAGCTGTGTGACGAGCTCC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCGACCCGCTACCTGGGGGCTTATACAT 1140
Db 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCGACCCGCTACCTGGGGGCTTATACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAAGGGGCCCATC 1200

DB 990 TGTGGAAATGTTCTGGATCCATTCCTTGTATGACAAACCTGGTTTACCCCTCAANTAG 1049
QY 601 GATGGAAGCTCTTGGCAAGATGGAAGAGATGTGGCCAGCAACCCCTGGTGAAC 660
DB 1050 GATGGAAGCTCTTGGCAAGATGGAAGAGATGTGGCCAGCAACCCCTGGTGAAC 1109
QY 661 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 1169
QY 721 GATGCGCTTATTCGATGATATGAGAGATGACAGCTTCTGTGAGTTCTCTGCCACAT 780
DB 1170 GATGCGCTTATTCGATGATATGAGAGATGACAGCTTCTGTGAGTTCTCTGCCACAT 1229
QY 781 GAGTGTGTGAACCAAGCCCGGACATATCTTGTGCTGCTGCGGCTGACAGGCTATCCCTGTG 840
DB 1230 GAGTGTGTGAACCAAGCCCGGACATATCTTGTGCTGCTGCGGCTGACAGGCTATCCCTGTG 1289
QY 841 GATGACAAACCGAAGCTGGCAAGATCAACGAATGTGAGCAGAGAACCAACAGCTGCAAC 900
DB 1290 GATGACAAACCGAAGCTGGCAAGATCAACGAATGTGAGCAGAGAACCAACAGCTGCAAC 1349
QY 901 CTGCAAGCAGACGCTGTCACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 960
DB 1350 CTGCAAGCAGACGCTGTCACAAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCCTGCT 1409
QY 961 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGATGCTGCTGCTGAGAACCCCTGCT 1020
DB 1410 GAGGAGCCTTATCTGAGATCACTGATTAACCGCTGATGCTGCTGCTGAGAACCCCTGCT 1469
QY 1021 TGCAGAGACCAAGCCCTTTACCATCTTGTACCGGAGCAATGAGAGCTGTGTGAGAGCTTC 1080
DB 1470 TGCAGAGACCAAGCCCTTTACCATCTTGTACCGGAGCAATGAGAGCTGTGTGAGAGCTTC 1529
QY 1081 GTTCCCGGCTGACATCTTCCAAATGCAAGCCAGCAACCCGCTAACCCCTGATTAACAT 1140
DB 1530 GTTCCCGGCTGACATCTTCCAAATGCAAGCCAGCAACCCGCTAACCCCTGATTAACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
DB 1590 TTCCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1649
QY 1201 AGTGCACCCCTGTGTATGACAGCCCATTAAGAGGCGCCGGAATTCAGCTGAGCTTG 1260
DB 1650 AGTGCACCCCTGTGTATGACAGCCCATTAAGAGGCGCCGGAATTCAGCTGAGCTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTCACTCACTTCAAGAGAGAGCTGCTGATCCGATCGGG 1320
DB 1710 GAAATGATCACTGTCAACACTGTCACTCACTTCAAGAGAGAGCTGCTGATCCGATCGGG 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 9
US-10-173-706-407
; Sequence 407, Application US/10173706
; Publication No. US2003002223A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
BEST FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-173-706-407

Query Match 99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCCAGGAATTAAGATATCTACATCTTACCATTCGCTGCTCTTCCAGCCCT 60
DB 450 ATGCCAGGAATTAAGATATCTACATCTTACCATTCGCTGCTCTTCCAGCCCT 509
QY 61 GGGAAATGACAGGACAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGACAGT 120
DB 510 GGGAAATGACAGGACAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGACAGT 569
QY 121 TTAGATATTTAGTAAATCCCGAATCCCGAGGCTCCCGAGAGACATGATGTGT 180
DB 570 TTAGATATTTAGTAAATCCCGAATCCCGAGGCTCCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGGGGTATTTATGATTCGCGGCAAAACCTGTATGAGAGGCGCTTC 240
DB 630 AACCAAAATGGGGGTATTTATGATTCGCGGCAAAACCTGTATGAGAGGCGCTTC 689
QY 241 TCGAACCCCTACTGACAGCCCTTACTAGTCCGTAAGCCAGCAGTGCACCACTCTCA 300
DB 690 TCGAACCCCTACTGACAGCCCTTACTAGTCCGTAAGCCAGCAGTGCACCACTCTCA 749
QY 301 GCTTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGATACAGATGAT 360
DB 750 GCTTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTGCACACCC 420
DB 810 GAAAGCAACCAATGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTGCACACCC 869
QY 421 ACCAGATCTGACATTAATCTGAAGGGGATACACCTCTCTGACACGAGGATAT 480
DB 870 ACCAGATCTGACATTAATCTGAAGGGGATACACCTCTCTGACACGAGGATAT 929
QY 481 CTTCGTGAGAGCCAGTCTTATGACATTTGATGATGTGCTTACTGCGACAGCTTC 540
DB 930 CTTCGTGAGAGCCAGTCTTATGACATTTGATGATGTGCTTACTGCGACAGCTTC 989
QY 541 TGTGCGAATGTTCTGATCTTATCTTGTATGATGCAACCCGTTTATACCTCAAT 600
DB 990 TGTGCGAATGTTCTGATCTTATCTTGTATGATGCAACCCGTTTATACCTCAAT 1049
QY 601 GATGGAAGCTCTTGGCAAGATGGAAGAGTGTGCAACCCCTGCGTGAAC 660
DB 1050 GATGGAAGCTCTTGGCAAGATGGAAGAGTGTGCAACCCCTGCGTGAAC 1109
QY 661 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGGCTCAACACCTAGAGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACCTTGAGAA 1169
QY 721 GATGCGCTTATTCGATGATATGAGAGATGACAGCTTCTGTGAGTTCTCTGCCACAT 780
DB 1170 GATGCGCTTATTCGATGATATGAGAGATGACAGCTTCTGTGAGTTCTCTGCCACAT 1229
QY 781 GAGTGTGTGAACCAAGCCCGGACATATCTTGTGCTGCTGCGGCTGACAGGCTATCCCTGTG 840
DB 1230 GAGTGTGTGAACCAAGCCCGGACATATCTTGTGCTGCTGCGGCTGACAGGCTATCCCTGTG 1289
QY 841 GATGACAAACCGAAGCTGGCAAGATCAACGAATGTGAGCAGAGAACCAACAGCTGCAAC 900

Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGCCAGGATAAAGGATCTACATGTTACATTCCTGGCTCTGCTCCAGCCCT 60
DB 450 ATGCCAGGATAAAGGATCTACATGTTACATTCCTGGCTCTGCTCCAGCCCT 509
QY 61 GGGAAATGCACAGGACAGTGCACGAATGGTTGACCTGATGTCGCACAGTGCAGT 120
DB 510 GGGAAATGCACAGGACAGTGCACGAATGGTTGACCTGATGTCGCACAGTGCAGT 569
QY 121 TTGATATTTGATTAATGCGCAACCATCCCGAGGCTCCGAGAGACATGATGTGT 180
DB 570 TTGATATTTGATTAATGCGCAACCATCCCGAGGCTCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGCGGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCTTAC 240
DB 630 AACCAAAATGGCGGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCTTAC 689
QY 241 TCGAACCCCTACTGACACCCCTACTAGTCCGATCCGATCCGAGAGCTGCCCACTCTCA 300
DB 690 TCGAACCCCTACTGACACCCCTACTAGTCCGATCCGATCCGAGAGCTGCCCACTCTCA 749
QY 301 GTCCTCAATATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 360
DB 750 GTCCTCAATATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCC 420
DB 810 GAAAGCAACCAATGTGTGATGTGAGAGATGTGCAACAGATTCACACAGTGCACACCC 869
QY 421 ACCCAGATCTGCATCAATCTGAAGCGGGGTACACCTGCTCTGTCACACCGAGATTTGG 480
DB 870 ACCCAGATCTGCATCAATCTGAAGCGGGGTACACCTGCTCTGTCACACCGAGATTTGG 929
QY 481 CTTTGGAGAGGCGGAGTGTAGACATGATGAATGGCTTGTGTTATGCCAGAGCTC 540
DB 930 CTTTGGAGAGGCGGAGTGTAGACATGATGAATGGCTTGTGTTATGCCAGAGCTC 989
QY 541 TGTGCAATGTCTCTGATCTCTATTTCTGTATACATGCAACCTGTGTTTACCCTCAATGAG 600
DB 990 TGTGCAATGTCTCTGATCTCTATTTCTGTATACATGCAACCTGTGTTTACCCTCAATGAG 1049
QY 601 GATGGAAGTCTTCCCAAGATGTGAACAGAGTGTGCCACCGAAGACCCCTGCGTGAACCC 660
DB 1050 GATGGAAGTCTTCCCAAGATGTGAACAGAGTGTGCCACCGAAGACCCCTGCGTGAACCC 1109
QY 661 TCGGTCAACACCTGAGGCTCTTCAATGCTGCGGTGTAACCCAGATATGAACTTGAGGA 720
DB 1110 TCGGTCAACACCTGAGGCTCTTCAATGCTGCGGTGTAACCCAGATATGAACTTGAGGA 1169
QY 721 GATGGCTTCATTTGATGATGATGAGAGTGCAGTTCCTGATGATCTCTGCCAAT 780
DB 1170 GATGGCTTCATTTGATGATGATGAGAGTGCAGTTCCTGATGATCTCTGCCAAT 1229
QY 781 GAGTGTGTACACAGCCCGGACATACCTTGTCTCTGCTCCAGGCTTACATCTCTGCTG 840
DB 1230 GAGTGTGTACACAGCCCGGACATACCTTGTCTCTGCTCCAGGCTTACATCTCTGCTG 1289
QY 841 GATGACAAACGGAAGTGTCCAAAGATCAAGGAATGTAGACACAGGAACCAACAGTGCAC 900
DB 1290 GATGACAAACGGAAGTGTCCAAAGATCAAGGAATGTAGACACAGGAACCAACAGTGCAC 1349
QY 901 CTGCAGACAGCTGTACATTTTACAAAGGGGGCTTCAATGATGACAGCCCATCCGCTGT 960
DB 1350 CTGCAGACAGCTGTACATTTTACAAAGGGGGCTTCAATGATGACAGCCCATCCGCTGT 1409
QY 961 GAGGAGCTTTATCTGAGATCACTGATTAACCGCTGTATGTGTCTGCTGAGAACCCTGGC 1020
DB 1410 GAGGAGCTTTATCTGAGATCACTGATTAACCGCTGTATGTGTCTGCTGAGAACCCTGGC 1469
QY 1021 TGCAGAGACAGCCCTTACCATCTTGTACCGGAGACATGAGAGCTGTACAGAGCTCC 1080
DB 1470 TGCAGAGACAGCCCTTACCATCTTGTACCGGAGACATGAGAGCTGTACAGAGCTCC 1529

```

```

QY 1081 GTTCCCGCTGACATCTTCCAAATGACACGACACCGCTACCTGGGCTATTACATT 1140
DB 1530 GTTCCCGCTGACATCTTCCAAATGACACGACACCGCTACCTGGGCTATTACATT 1589
QY 1141 TTCCAGATCAATCTGGGAATGAGGCGAGAGATTTTACATCGGGCAACGGGCCCCATC 1200
DB 1590 TTCCAGATCAATCTGGGAATGAGGCGAGAGATTTTACATCGGGCAACGGGCCCCATC 1649
QY 1201 AGTGCACCCCTGGTGTATGACACGCCCCATCAAAAGGCCCCGGGAAATCCAGCTGGACTTG 1260
DB 1650 AGTGCACCCCTGGTGTATGACACGCCCCATCAAAAGGCCCCGGGAAATCCAGCTGGACTTG 1709
QY 1261 GAAATGATCACTGTCAACACATGTCATCAACTTCAGAGGCGAGCTCCGTGATCCGACTGCG 1320
DB 1710 GAAATGATCACTGTCAACACATGTCATCAACTTCAGAGGCGAGCTCCGTGATCCGACTGCG 1769
QY 1321 ATATATGTGTGCGAGTACCCATTTC 1344
DB 1770 ATATATGTGTGCGAGTACCCATTTC 1793

```

RESULT 7

```

US-10-176-758-407
; Sequence 407, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-758-407

```

Query Match

99.9%: Score 1342.4; DB 9; Length 2609;

Best Local Similarity 99.9%: Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

```

Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCAGGATAAAGGATCTACATGTTACATTCCTGGCTCTGCTCCAGCCCT 60
DB 450 ATGCCAGGATAAAGGATCTACATGTTACATTCCTGGCTCTGCTCCAGCCCT 509
QY 61 GGGAAATGCACAGGACAGTGCACGAATGGTTGACCTGATGTCGCACAGTGCAGT 120
DB 510 GGGAAATGCACAGGACAGTGCACGAATGGTTGACCTGATGTCGCACAGTGCAGT 569
QY 121 TTGATATTTGATTAATGCGCAACCATCCCGAGGCTCCGAGAGACATGATGTGT 180
DB 570 TTGATATTTGATTAATGCGCAACCATCCCGAGGCTCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGCGGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCTTAC 240
DB 630 AACCAAAATGGCGGGATTTATGATCTCCCGGACAAACCTGTGTATCGAGGCTTAC 689
QY 241 TCGAACCCCTACTGACACCCCTACTAGTCCGATCCGATCCGAGAGCTGCCCACTCTCA 300

```



```

; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: PCT/US98/25190
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

```

```

Query Match      99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGCCGAGATTAAGAGATCTACTGTTACCATTTCTGCTCTCTTCCAAAGCCCT 60
DB 450 ATGCCGAGATTAAGAGATCTACTGTTACCATTTCTGCTCTCTTCCAAAGCCCT 509
QY 61 GGGAAATGACAGAGCAGACAGTSCAGAAATGGCTTTGACCTGGATGGCCAGTCAGAGACAGTGT 120
DB 510 GGGAAATGACAGAGCAGACAGTSCAGAAATGGCTTTGACCTGGATGGCCAGTCAGAGACAGTGT 569
QY 121 TTGATATTTGATGATATGATCCGACCAATCCCGAGAGCCCTGCGAGAGAGACATGATGTGT 180
DB 570 TTGATATTTGATGATATGATCCGACCAATCCCGAGAGCCCTGCGAGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGGGGATTTATGATTTCCCGGAGAAACCCCTGTGTATCGAGGGCCCTTAC 240
DB 630 AACCAAAATGGGGGATTTATGATTTCCCGGAGAAACCCCTGTGTATCGAGGGCCCTTAC 689
QY 241 TCGAACCCCTACTCGACCCCTACTCAGTCCGTACCCAGACAGTGGCCCAACACTCTCA 300
DB 690 TCGAACCCCTACTCGACCCCTACTCAGTCCGTACCCAGACAGTGGCCCAACACTCTCA 749
QY 301 GCTCCAAATCTATCCCAAGATCTCCAGGCTCTTATATGCGCGTTTGGATACCAAGATGAT 360
DB 750 GCTCCAAATCTATCCCAAGATCTCCAGGCTCTTATATGCGCGTTTGGATACCAAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTTCCCAACCAAGATGAT 420
DB 810 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTTCCCAACCAAGATGAT 869
QY 421 ACCGATCTGCATCAATAGTAAGGCGGGTACACCTGCTCTGCAACGAGGATATTGG 480
DB 870 ACCGATCTGCATCAATAGTAAGGCGGGTACACCTGCTCTGCAACGAGGATATTGG 929
QY 481 CTTCTGGAAGGCAATGCTTACACATGTATGATGTGCTGTATGCTTACTGCCAGACCTC 540
DB 930 CTTCTGGAAGGCAATGCTTACACATGTATGATGTGCTGTATGCTTACTGCCAGACCTC 989
QY 541 TGTGGAATGTTCTCGATATCTATCTTATACATGCAACCGTGTATACCCCAATGAG 600
DB 990 TGTGGAATGTTCTCGATATCTATCTTATACATGCAACCGTGTATACCCCAATGAG 1049
QY 601 GATGGAAGGCTTGGCAAGATGTGAACAGAGTGTGCCACCGAGAACCCCTGCTGCAAAAC 660
DB 1050 GATGGAAGGCTTGGCAAGATGTGAACAGAGTGTGCCACCGAGAACCCCTGCTGCAAAAC 1109
QY 661 TGGGTCAACACCTAGGCTCTTTCATCTGCGCGCTGTGACCCGAGATATGAATTTGAGAA 720
DB 1110 TGGGTCAACACCTAGGCTCTTTCATCTGCGCGCTGTGACCCGAGATATGAATTTGAGAA 1169
QY 721 GATGGCGTTTCAATGAGTATGAGAGAGTGCACCTTCTGAGTGTCTCTGCCAAACAT 780
DB 1170 GATGGCGTTTCAATGAGTATGAGAGAGTGCACCTTCTGAGTGTCTCTGCCAAACAT 1229

```

```

QY 781 GAGTGTGTGAACACAGCCCGGACATACTTCTGCTCTGCTCCCTCCAGGCTACATCTGCTG 840
DB 1230 GAGTGTGTGAACACAGCCCGGACATACTTCTGCTCTGCTCCCTCCAGGCTACATCTGCTG 1289
QY 841 GATGACAAACCGAAGCTTCCCAAGACATCAACGAATGTGAGCAGACAGAACCAACGTCGAAC 900
DB 1290 GATGACAAACCGAAGCTTCCCAAGACATCAACGAATGTGAGCAGACAGAACCAACGTCGAAC 1349
QY 901 CTCGACGACAGTGTCTCAATTTTCAAGGGGGCTTCAAGTGCATGACACCCATCCGCTGT 960
DB 1350 CTCGACGACAGTGTCTCAATTTTCAAGGGGGCTTCAAGTGCATGACACCCATCCGCTGT 1409
QY 961 GAGGACCTTATCTGAGATGAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1020
DB 1410 GAGGACCTTATCTGAGATGAGTATACCGCTGTATGTCTCTGCTGAGAACCTTGGC 1469
QY 1021 TGCAGAGACACAGCCCTTTACCATCTTGTACCGGGACATGACAGTGTGTCAAGAGCTTC 1080
DB 1470 TGCAGAGACACAGCCCTTTACCATCTTGTACCGGGACATGACAGTGTGTCAAGAGCTTC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTACAT 1140
DB 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGGCTATTACAT 1589
QY 1141 TTCCAGATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGCGGCCCATC 1200
DB 1590 TTCCAGATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGCGGCCCATC 1649
QY 1201 AGTGCACCCCTGGTATGACACAGCCCATAAAGGGGCGGGGAAATCCAGCTGGACCTG 1260
DB 1650 AGTGCACCCCTGGTATGACACAGCCCATAAAGGGGCGGGGAAATCCAGCTGGACCTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGCAGCTCCGTATCCGATCGAGT 1320
DB 1710 GAAATGATCACTGTCAACACTGTCAATCACTTCAAGAGCAGCTCCGTATCCGATCGAGT 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1770 ATATATGTGTGCGAGTACCATTC 1793

```

RESULT 6

US-10-174-590-407

; Sequence 407, Application US/10174590

; Publication No. US20030008352A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C42

; CURRENT APPLICATION NUMBER: US/10-174, 590

; PRIOR FILING DATE: 2002-06-18

; PRIOR APPLICATION REMOVED - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 407

; LENGTH: 2609

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-174-590-407

```

Query Match      99.9%; Score 1342.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;

```


PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139655
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804

PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17

LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORN10701
CLONE: 45517
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-836-561-2

Query Match 100.0%; Score 1344; DB 10; Length 2550;
Best Local Similarity 100.0%; Pident. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCCAGGAAATTAAGGATATCTACCTTACCATTTCTGGGCTCTGCTTCCAGCCCT 60
382 ATCCAGGAAATTAAGGATATCTACCTTACCATTTCTGGGCTCTGCTTCCAGCCCT 441
61 GGGAAATGCACAGGCACAGTGCACGAAATGGCTTTGACCTGGATCCGACATCAGAGAGTGT 120
442 GGGAAATGCACAGGCACAGTGCACGAAATGGCTTTGACCTGGATCCGACATCAGAGAGTGT 501
121 TTGATATTTGATGATATCCGAAATCCCGAGGCTTCCGAGAGACATGATGTGT 180
502 TTGATATTTGATGATATCCGAAATCCCGAGGCTTCCGAGAGACATGATGTGT 561
181 AACCAAAATGGGGGTTATTTATGCTATCCCGGACAAACCTGTGTATCGAGGGGCTTAC 240
562 AACCAAAATGGGGGTTATTTATGCTATCCCGGACAAACCTGTGTATCGAGGGGCTTAC 621
241 TCGAACCCCTACTCGACCCCTACTCAGATCCGATCCGAGAGAGTGCACACACTCTCA 300
622 TCGAACCCCTACTCGACCCCTACTCAGATCCGATCCGAGAGAGTGCACACACTCTCA 681
301 GCTCCAAATCTCCAGATCTCGAGGCTTTATATGCCGCTTTGGATTACCATGTGAT 360
682 GCTCCAAATCTCCAGATCTCGAGGCTTTATATGCCGCTTTGGATTACCATGTGAT 741
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACAGATTTCCACACAGTGCACACCC 420
742 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACAGATTTCCACACAGTGCACACCC 801
421 ACCCAATCTGATCAATACTGAAAGGGGGGTACACCTGCTCTGACCGAGGATATTGG 480
802 ACCCAATCTGATCAATACTGAAAGGGGGGTACACCTGCTCTGACCGAGGATATTGG 861
481 CTTCTGGAAGGCCAGTGTAGACATGATGATGTGCTTGTGTTACTGCGACAGCTC 540
862 CTTCTGGAAGGCCAGTGTAGACATGATGATGTGCTTGTGTTACTGCGACAGCTC 921
541 TGTGCAATGTCTGATCTTATTTGTATGATGCAACCTGTGTTTACCTCAATGAG 600
922 TGTGCAATGTCTGATCTTATTTGTATGATGCAACCTGTGTTTACCTCAATGAG 981
601 GATGAGAGGTCTTGCCAGATGTGAAAGAGTGTGCCACCGAAGCCCTGGGTGCAACC 660
982 GATGAGAGGTCTTGCCAGATGTGAAAGAGTGTGCCACCGAAGCCCTGGGTGCAACC 1041
661 TGGGTAAACACCTTACGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTAGGAA 720
1042 TGGGTAAACACCTTACGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTAGGAA 1101
721 GATGGGCTTATGAGTATGAGAGAGTGCAGGCTTCTGATGCTCTGCGCAACAT 780
1102 GATGGGCTTATGAGTATGAGAGAGTGCAGGCTTCTGATGCTCTGCGCAACAT 1161
781 GAGTGTGTGAACACCGCCGACATATCTGCTCTGCGCTCCAGGCTTACATCTGTGTG 840
1162 GAGTGTGTGAACACCGCCGACATATCTGCTCTGCGCTCCAGGCTTACATCTGTGTG 1221
841 GATGACACCGGAGAGTGCAGAGATCAAGATGTGAGACACAGAGACACACGTGCAC 900
1222 GATGACACCGGAGAGTGCAGAGATCAAGATGTGAGACACAGAGACACACGTGCAC 1281

QY 901 CTGACAGACGCTGTCAATTTTACAGGGGGCTTCAATGATGACACCCATCCGCTGT 960
DB 1282 CTGACAGACGCTGTCAATTTTACAGGGGGCTTCAATGATGATGACACCCATCCGCTGT 1341
QY 961 GAGAGGCTTATCTGAGATGATGATTAACCGCTGTATGTCTGCTGAGAAACCTGTGC 1020
DB 1342 GAGAGGCTTATCTGAGATGATGATTAACCGCTGTATGTCTGCTGAGAAACCTGTGC 1401
QY 1021 TGCAGAGACACGCTTTTACATCTTGTACCGGGACATGAGAGTGTGTCAAGAGCTTC 1080
DB 1402 TGCAGAGACACGCTTTTACATCTTGTACCGGGACATGAGAGTGTGTCAAGAGCTTC 1461
QY 1081 GTTCCGCTGACATCTTCAATGACAGCCGCTTACCGGGGCTTATGATGAT 1140
DB 1462 GTTCCGCTGACATCTTCAATGACAGCCGCTTACCGGGGCTTATGATGAT 1521
QY 1141 TTCCAGATCAATTTGGAATGAGGGGAGAGATTTTACATGCGGCAAAACGGGCCCCATC 1200
DB 1522 TTCCAGATCAATTTGGAATGAGGGGAGAGATTTTACATGCGGCAAAACGGGCCCCATC 1581
QY 1201 AGTGCACCCCTGCTGATGACACGCCCATCAAGGGCCCCGGGAAATCCAGCTGACTTG 1260
DB 1582 AGTGCACCCCTGCTGATGACACGCCCATCAAGGGCCCCGGGAAATCCAGCTGACTTG 1641
QY 1261 GAATGATCACTGTCAACACTGTATCAACTTCAAGAGGACGCTCCGATCCGACTGGGG 1320
DB 1642 GAATGATCACTGTCAACACTGTATCAACTTCAAGAGGACGCTCCGATCCGACTGGGG 1701
QY 1321 ATATATGTGTGCGAGTACCCATTC 1344
DB 1702 ATATATGTGTGCGAGTACCCATTC 1725

RESULT 5

US-10-066-500-14
Sequence 14, Application US/10066500
Patent No. US2002017165A1
GENERAL INFORMATION:
APPLICANT: Avi J. Ashkenazi
APPLICANT: Kevin P. Baker
APPLICANT: David A. Bolstein
APPLICANT: Luc Desnoyers
APPLICANT: Dan L. Eaton
APPLICANT: Napoleone Ferrara
APPLICANT: Sherman Fong
APPLICANT: Wei-Qiang Gao
APPLICANT: Hanspeter Gerber
APPLICANT: Mary E. Gerltisen
APPLICANT: Audrey Goddard
APPLICANT: Paul J. Godowski
APPLICANT: Austin L. Gurney
APPLICANT: Ivar J. Kljavin
APPLICANT: Jennie P. Mather
APPLICANT: Mary A. Napier
APPLICANT: James Pan
APPLICANT: Nicholas F. Paoni
APPLICANT: Margaret Ann Roy
APPLICANT: Timothy A. Stewart
APPLICANT: Daniel Tumas
APPLICANT: Colin K. Watanabe
APPLICANT: P. Mickey Williams
APPLICANT: William I. Wood
APPLICANT: Zemin Zang
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P130R1C7
CURRENT APPLICATION NUMBER: US/10/066,500
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 10/002,796
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26

SEQUENCE CHARACTERISTICS:
 LENGTH: 2362 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-041-016-1

Query Match 100.0%; Score 1344; DB 9; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; gaps 0;

```

QY 1 ATGCCGAGATTAAGAGATCTACATCTGTTACCATTTGCGCTCTGCTCTTCCAAAGCCCT 60
DB 186 ATGCCGAGATTAAGAGATCTACATCTGTTACCATTTGCGCTCTGCTCTTCCAAAGCCCT 245
QY 61 GGGATGACACAGGCACAGTGCAGCAATGGCTTTGACCTGGATGCGCAGTCAGAGACAGTGT 120
DB 246 GGGATGACACAGGCACAGTGCAGCAATGGCTTTGACCTGGATGCGCAGTCAGAGACAGTGT 305
QY 121 TTAGATATTGATTAATGCCGACACCATCCCGAGGCGCTGCGAGAGACATGATGTGT 180
DB 306 TTAGATATTGATTAATGCCGACACCATCCCGAGGCGCTGCGAGAGACATGATGTGT 365
QY 181 AACCAAAATGGCGGGATTATTATGATTCGCCGACAAACCGCTGTATCGAGGGCCCTAC 240
DB 366 AACCAAAATGGCGGGATTATTATGATTCGCCGACAAACCGCTGTATCGAGGGCCCTAC 425
QY 241 TCGAACCCCTACTCGAACCCCTACTCAGATCCGTACCCAGACAGCTGCCACCACTCTCA 300
DB 426 TCGAACCCCTACTCGAACCCCTACTCAGATCCGTACCCAGACAGCTGCCACCACTCTCA 485
QY 301 GCTCCAAATATCTCCAGATCTCCAGGCGCTTATATGCCGCTTGTGATACACATGAT 360
DB 486 GCTCCAAATATCTCCAGATCTCCAGGCGCTTATATGCCGCTTGTGATACACATGAT 545
QY 361 GAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCCAACCGATGCAACCC 420
DB 546 GAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCCCAACCGATGCAACCC 605
QY 421 ACCGAGTCTGATCAATATCTGAAAGCGGGGTACACCTGCTCTGACACCGAGATATTTGG 480
DB 606 ACCGAGTCTGATCAATATCTGAAAGCGGGGTACACCTGCTCTGACACCGAGATATTTGG 665
QY 481 CTTTGGAAAGGCGGAGGCTTAGCATGTGATGTGCGTATGCTTACTGCGACAGCTC 540
DB 666 CTTTGGAAAGGCGGAGGCTTAGCATGTGATGTGCGTATGCTTACTGCGACAGCTC 725
QY 541 TGTGGGAATGTTCTGTGATCTTATCTTGTACATGCAACCGTGGTTTACCTCAATGAG 600
DB 726 TGTGGGAATGTTCTGTGATCTTATCTTGTACATGCAACCGTGGTTTACCTCAATGAG 785
QY 601 GATGGAAGTCTTCCCAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGTCAAAAC 660
DB 786 GATGGAAGTCTTCCCAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGTCAAAAC 845
QY 661 TGGCGTCAACCTAGCGGCTTTTCATCTGCGCTGTGACCCAGATATGAATTTAGAGAA 720
DB 846 TGGCGTCAACCTAGCGGCTTTTCATCTGCGCTGTGACCCAGATATGAATTTAGAGAA 905
QY 721 GATGGCGTTTCATCTGAGTATGTGACAGAGTCAAGTCTTCTGAGTTCTCTGCCAAT 780
DB 906 GATGGCGTTTCATCTGAGTATGTGACAGAGTCAAGTCTTCTGAGTTCTCTGCCAAT 965
QY 781 GAGTGTGGAACACCGCGGACATCTTCTGCTCTGCGCTTCCAGAGCTACATCTGCTG 840
DB 966 GAGTGTGGAACACCGCGGACATCTTCTGCTCTGCGCTTCCAGAGCTACATCTGCTG 1025
QY 841 GATGACAACCGAGGTGCGCAAGACATCAAGAAATGTGACACAGAAACACAGCTGCAAC 900
DB 1026 GATGACAACCGAGGTGCGCAAGACATCAAGAAATGTGACACAGAAACACAGCTGCAAC 1085

```

```

QY 901 CTGCAGACAGCTGTCTACAAATTTCAAGGGGGCTTCAATGATCGACCCCATCGCTGT 960
DB 1086 CTGCAGACAGAGTGTCTCAATTTTCAAGGGGGCTTCAATGATCGACCCCATCGCTGT 1145
QY 961 GAGGAGCCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGGC 1020
DB 1146 GAGGAGCCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCCCTGGC 1205
QY 1021 TGCAGAGACAGCCCTTACATCTTGTACCGGGACATGAGACGTGTCTGAGAGCTCC 1080
DB 1206 TGCAGAGACAGCCCTTACATCTTGTACCGGGACATGAGAGCTGTCTGAGAGCTCC 1265
QY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCTGGGGCTTATTAAT 1140
DB 1266 GTTCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCTGGGGCTTATTAAT 1325
QY 1141 TTCCAGATCAAAATCTGGGATGAGGGGAGAGAAATTTACATGCGGCAAAAGCCCATTC 1200
DB 1326 TTCCAGATCAAAATCTGGGATGAGGGGAGAGAAATTTACATGCGGCAAAAGCCCATTC 1385
QY 1201 AGTCCACCCCTGTGTATGACACGCCCCATCAAGGGGCCCGGAAATTCAGCTGACTTG 1260
DB 1386 AGTCCACCCCTGTGTATGACACGCCCCATCAAGGGGCCCGGAAATTCAGCTGACTTG 1445
QY 1261 GAATGATCACTGTCAACACTGTCACTTCACTTCAAGAGGACGCTCGTATCCGACTGGCG 1320
DB 1446 GAATGATCACTGTCAACACTGTCACTTCACTTCAAGAGGACGCTCGTATCCGACTGGCG 1505
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1506 ATATATGTGTGCGAGTACCATTC 1529

```

RESULT 3

US-09-083-002-1

Sequence 1, Application US/09083002

Patent No. US2001001650A1

GENERAL INFORMATION:

```

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racie, Lisa A.
APPLICANT: Lavallee, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Agostino, Michael
APPLICANT: Lu, Zhijian
APPLICANT: Honjo, Tasuku
APPLICANT: Tashiro, Kei
APPLICANT: Nakamura, Tomoyuki
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284

```

MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 211..1554
 US-09-275-805-1

Query Match 100.0%; Score 1344; DB 10; Length 1717;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCCGGAATATAAAGATACACCTGTTACCATTTGCTCTCTGCTTCCAAAGCCCT 60
DB 211 ATGCCGGAATATAAAGATACACCTGTTACCATTTGCTCTCTGCTTCCAAAGCCCT 270
QY 61 GGAATGACAGCAGCAGTGCAGATGGCTTTGACCGGATCGGACGACGAGAGTGT 120
DB 271 GGAATGACAGCAGCAGTGCAGATGGCTTTGACCGGATCGGACGAGAGTGT 330
QY 121 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 331 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
QY 181 AACCAAAAGGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 391 AACCAAAAGGGGGGATTTATGATGATGATGATGATGATGATGATGATGATGAT 450
QY 241 TCGAACCCCTACTGACCCCTACTGACGTCCTGACCGAGAGCTGCCCACTCTCA 300
DB 451 TCGAACCCCTACTGACCCCTACTGACGTCCTGACCGAGAGCTGCCCACTCTCA 510
QY 301 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGATATCCAGAT 360
DB 511 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTTGATATCCAGAT 570
QY 361 GAAAGAACCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 420
DB 571 GAAAGAACCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 630
QY 421 ACCCAGATCTGCATCATCTGATGAGGCGGTACACCTGCTCTGACCGAGAGATTTG 480
DB 631 ACCCAGATCTGCATCATCTGATGAGGCGGTACACCTGCTCTGACCGAGAGATTTG 690
QY 481 CTTCTGGAAGGCGAGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 540
DB 691 CTTCTGGAAGGCGAGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 750
QY 541 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 751 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810
QY 601 GATGGAAGTCTTGCCAAAGATGTGAAGAGTGTGACCGAGAACCCCTGCTGCAAAAC 660
DB 811 GATGGAAGTCTTGCCAAAGATGTGAAGAGTGTGACCGAGAACCCCTGCTGCAAAAC 870
QY 661 TGGGTCAACACTGACGCTCTTTCATCTGCGGTGTGACCGAGATATGAACCTTGA 720
DB 871 TGGGTCAACACTGACGCTCTTTCATCTGCGGTGTGACCGAGATATGAACCTTGA 930
QY 721 GATGGGCTTCATGCACTGATGATGAGAGTGTGACGCTCTGAGTCTCTGCAACAT 780
DB 931 GATGGGCTTCATGCACTGATGATGAGAGTGTGACGCTCTGAGTCTCTGCAACAT 990
QY 781 GATGTGTGAACAGCCCGGACATCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCT 840
DB 991 GATGTGTGAACAGCCCGGACATCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCT 1050
QY 841 GATGACAAACGAAAGTGTGCAAGACATCAACGAATGTGAGACAGGAAACCAAGTGAAC 900
DB 1051 GATGACAAACGAAAGTGTGCAAGACATCAACGAATGTGAGACAGGAAACCAAGTGAAC 1110
QY 901 CTGCACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1111 CTGCACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170

```

```

QY 961 GAGAGCCCTATCTGAGGATCAGTATACCGCTGTATGTGCTGTGAGAACCCCTGAC 1020
DB 1171 GAGAGCCCTATCTGAGGATCAGTATACCGCTGTATGTGCTGTGAGAACCCCTGAC 1230
QY 1021 TGCAGAGACAGCCCTTTACCATCTTGTACCGGACATGACAGTGTGTACAGAGCTCC 1080
DB 1231 TGCAGAGACAGCCCTTTACCATCTTGTACCGGACATGACAGTGTGTACAGAGCTCC 1290
QY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTCCCTGCGGCTTATACAT 1140
DB 1291 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTCCCTGCGGCTTATACAT 1350
QY 1141 TTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTACATGCGGCAAGCGGCCCATC 1200
DB 1351 TTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTACATGCGGCAAGCGGCCCATC 1410
QY 1201 AGTGCACCCCTGTGTGATGACAGCCCATCAAAAGGCGCCCGGAAATCCAGCTGACTTG 1260
DB 1411 AGTGCACCCCTGTGTGATGACAGCCCATCAAAAGGCGCCCGGAAATCCAGCTGACTTG 1470
QY 1261 GAAATGATCACTGCAACACTGTATCACTTACAGAGGAGCTCGTGTATCCGACTGCG 1320
DB 1471 GAAATGATCACTGCAACACTGTATCACTTACAGAGGAGCTCGTGTATCCGACTGCG 1530
QY 1321 ATATATGTGTGCGAGTACCCATTTC 1344
DB 1531 ATATATGTGTGCGAGTACCCATTTC 1554

```

RESULT 2 US-10-041-016-1

Sequence 1, Application US/10041016

Patent No. US20020165151A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.

Racie, Lisa A.

Lavallee, Edward R.

Merberg, David

Treacy, Maurice

Evans, Cheryl

Agostino, Michael

Lu, Zhijian

Honjo, Tasuku

TITLE OF INVENTION: SECRETED PROTEINS

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/041,016

FILING DATE: 07-Jan-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/083,002

FILING DATE: 21-MAR-1998

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.

REGISTRATION NUMBER: P-41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 1:

Sat Jul 5 15:31:05 2003

us-09-674-379a-11.rge

Page 16.

QY	1297	GGCAGCTCCGTGATCCGACTGCCGATATATGTCGCAGTAACCATTC	1344
Dd	1378	GCCAGCTCTGTACTGAGGCTCACCGTCTTTGTAGGGGCCATACACCTTC	1425

Search completed: July 3, 2003, 15:34:23
Job time : 3509.57 secs

QY 937 AATGATGACCCCATCGCTGTAGAGACCTTATCTGAGATCAGTGAATACCGCTGT 996
Db 922 CGCTGCGTGGACACCAACCGCTGTGAGCCCTACATCCAGTCTCTGAGAACCCCTGT 981
QY 997 ATGTGTCTGTGAGAACCTGGCTGACAGAGACACCGCTTACCACTTGTACCGGGAC 1056
Db 982 CTGTGCGCGGCTCCACACCTCTATGTGAGAGACAGCCCTTACCTTTGTGACCGCTAC 1041
QY 1057 ATGACGCTGTGTAGAGACCTCCCTGTCGCTGACATCTTCCAAATGCAAGCCAGAC 1116
Db 1042 ATGACCAATCACCCTGGAGACGCTGCGCTGACATCTTCCAAATGCAAGCCAGAC 1101
QY 1117 CGTACCCCTGGGCTTATTAATTTTCCAGATCAATTTGGAAATGAGGCGAGAAATTT 1176
Db 1102 GTTACACCCGCTGCTACAAATGCTTTTACATCTGCTGGAATCTCGAGGGGAGACTTT 1161
QY 1177 TACTGTGGGCAAGAGGGGCCCATGATGATGATGATGATGATGATGATGATGATGATG 1236
b 1162 TACTTATGAGCAATCAACACGATGAGCGCATGCTGCTGCTGCTGCTGCTGCTGCTG 1221
QY 1237 CCGCGGGAATTCAGCTGAGTGGAAATGATGATGATGATGATGATGATGATGATGATG 1296
Db 1222 CCGCGGGAATTCAGCTGAGTGGAAATGATGATGATGATGATGATGATGATGATGATG 1281
QY 1297 GCGAGCTCGGTGATCCGATGCGGATATATGTGTGAGTACATCCATTC 1344
Db 1282 GCCAGCTGTACTGAGAGCTACCGCTCTTTTGTAGGGGCTACACCTTC 1329

RESULT 15
AB030655 1536 bp mRNA linear PRI 25-MAR-2000
LOCUS Homo sapiens mRNA for mutant p53 binding protein 1 (MBP1), complete
DEFINITION
AB030655
VERSION MBP1, mutant p53 binding protein 1 (MBP1).
KEYWORDS Homo sapiens cDNA to mRNA.
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites) Tanaka, S., Sugimachi, K. and Sugimachi, K.
TITLE Human mutant p53 binding protein (MBP1)
JOURNAL Unpublished
AUTHORS Tanaka, S.
2 (bases 1 to 1536)
REFERENCE
1 (sites) Tanaka, S.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1999) Shinji Tanaka, Kyushu University, Faculty of
Medicine, Department of Surgery II, 3-1-1 Maidashi, Higashi-ku,
Fukuoka 812-8582, Japan (E-mail: shinji@sur2.med.kyushu-u.ac.jp,
Tel: 81-92-642-5466, Fax: 81-92-642-5482)
LOCATION/Qualifiers
1. 1536
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 1536
/gene="MBP1"
97. 1428
/note="Human homologue of murine mutant p53 binding
protein 1 (MBP1) reported by Gallagher et al. (Oncogene
18; 3608-16, 1999)."
/codon_start=1
/product="mutant p53 binding protein 1 (MBP1)."
/protein_id="BA092880.1"
/db_xref="GI:7328921"

FEATURES
source
gene
CDS

ATRLCQDIDECESGAHOCSEAOCTVNEHGGY RCYDPTNRCVPEYIOVSENRCLCPASNP
LCREOPSIYHYMTITSERVAPADVDIOIA.TSVYRGAN.FOITAGSOGDFYTRQI
NNVSAMVILAPVTPREHYDLEMTVNSIMLSYRASSVILNLTFFVGAATP"

BASE COUNT 307 a 500 c 436 g 293 t

ORIGIN

Query Match 27.7%; Score 372.8; DB 9; Length 1536;
Best Local Similarity 61.3%; Pred. No. 9.2e-97;
Matches 618; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

QY 337 TCGCGCTTTGGATACCATGATGATAAAGCAATGTTGATGATGAGAGTGTCA 396
Db 421 TGGCCACAGCATATGAGCCGAGAGATGAGAGACAGTGTGTGATGATGAGAGTGTCC 480
QY 397 ACAGATTCCACAGTACCAACCCAGCAGTGTGATGATGATGATGATGATGATGATGATG 456
Db 481 CAGGCCCTGACACACATCTGCGCCAGCAGAGACTGCGCATTAATCTTGCCTGCTATCAG 540
QY 457 TGTCTGTGACGACGAGATATTTGCTTGGAAAGCCAGTGTGATGATGATGATGATGATG 516
Db 541 TGCACCTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 517 CGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
Db 601 CGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 577 AACCTGTGTTTACCTCAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 636
Db 661 GAGCGGGTTTCCAGTGTGGGGCTTAACACCCGCTCTGTTGTATGATGATGATGATGATG 720
QY 637 ACCGAGAACCCCTGCGTCAAAACCTGCTGATCAACACCTGCTTTCATCTGCGCTGT 696
Db 721 ATGGGGGCCCATGTCAGAGAGCGCTCTTCAACCTCATGATGATGATGATGATGATGATG 780
QY 697 GACCAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
Db 781 CACAGAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 757 TTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 816
Db 841 TACTCAGCTACCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 817 TGCCTTCAGGCTATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 876
Db 901 TGCCACAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 957
QY 877 GAGCAGAGAACACACAGTGAACCTGACAGAGTGTGATGATGATGATGATGATGATGATG 936
Db 958 GAGTGTGTTGGCCACACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1017
QY 937 AATGATGACCCCATCGCTGTGAGAGACCTTATCTGAGATCAGTGAATACCGCTGT 996
Db 1018 CGTGTGCGGAGACACCAACCGCTGTGAGAGCTTATCTGAGATCAGTGAATACCGCTGT 1077
QY 997 ATGTGTCTGTGAGAACCTGGCTGACAGAGACCGCTTATCTGAGATCAGTGAATACCGCTGT 1056
Db 1078 CTGTGCGCGGCTCCACACCTCTATGTGAGAGAGCTTATCTGAGATCAGTGAATACCGCTGT 1137
QY 1057 ATGACGCTGTGTAGAGAGCTCCCTGTCGCTGACATCTTCCAAATGCAAGCCAGAC 1116
Db 1138 ATGACCAATCACCCTGGAGACGCTGCGCTGACATCTTCCAAATGCAAGCCAGAC 1197
QY 1117 CGTACCCCTGGGCTTATTAATTTTCCAGATCAATTTGGAAATGAGGCGAGAAATTT 1176
Db 1198 GTTACACCCGCTGCTACAAATGCTTTCAGATCTGCTGGAATCTCCAGAGGGAGACTTT 1257
QY 1177 TACTGTGGGCAAGAGGGGCCCATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
Db 1258 TACTTATGAGCAATCAACACGATGAGCGCATGCTGTGCTGCTGCTGCTGCTGCTGCTG 1317
QY 1237 CCGCGGGAATTCAGCTGAGTGTGAAATGATGATGATGATGATGATGATGATGATGATG 1296
Db 1318 CCGCGGGAATTCAGCTGAGTGTGAAATGATGATGATGATGATGATGATGATGATGATGATG 1377

OY		661	TGGGTCAACACCAGGCGCTTTTCAATCCGCCCTGTGACCCAGATATGAATTGAGAA	720
Db		1010	TGGGTCAACACTATGGTTCTTTCATCTCCCTGTGACCCAGATATGAATTGAGAA	1061
OY		721	GATGGCGTTTCATTGCAGTATATGACGAGTGACGTTCTCGAGTTCCCTGCCAAT	780
Db		1070	GATGGCATTCACATCGCAGTGATATGGATGATGAGTGACGTTCTCGAGTTCTCTGCACAT	1122
OY		781	GAGTGTGTAACCAAGCCCCGGCACATTAATCTCTGCTCTGCTCCACAGCTACT	840
Db		1130	GAGTGTGTAAACCAAGCCCCGGGCTCAATCTTCTCATATGCCCCTCCAGGCTACTCTTGTG	1189
OY		841	GATGCAACACGGAAGCTGCCAACAACATCAAGAATGAGACACAGAACCAACAGCGCAAC	900
Db		1190	GAAGATAACCGAAGCTGCCAGGATATCAATGAATGTGACACCGGAACCAACATGACACT	1243
OY		901	CTGCACAGACGCTGCTACATTTTACAAAGGGGGCTTCAAATGCATGACCCCATCCGCTGT	960
Db		1250	CCCCCTGCACACTTGTCTACATTTGCAAGGGGGCTTCAAATGATGACCCCATCTGTGCG	1300
OY		961	GAGGAGCCCTTATCTGAGATCAAGTATAACCGCTGATGTGCTGTGCTGAGAACCCCTGGC	1020
Db		1310	GAGGAGCCCTTATCTGAGATCAAGTATAACCGCTGATGTGCTGTGCTGAGAACCCCTGGC	1365
OY		1021	TGCAGAGACACCCCTTTACACTCTTGTACCGGGACATGAGACGTGTGTCAGACGCTCC	1080
Db		1370	TGCAGAGAGACACCCCATTCACACTCTTGTGTGGGACATGAGATGTGTATCAGAACGCTCT	1429
OY		1081	GTYCCGGCGATCTTCCAAATGGCAACGACACCCGCTACCTCGGGGCTTATTAATCT	1140
Db		1430	GTYCCGCTGACATTTTCCAGATGCAACGACACCCGATATCCCTGCGGCTTATTAATCT	1489
OY		1141	TTCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGGGCCCATC	1200
Db		1490	TTCAGATCAAAATCTGGGAACGAGGGGTGAGAGTTCTACATGCGGCAAAACGGGCGTATC	1549
OY		1201	AGTGCACCCCTGGTGATGATACAGACCCCATTCAAAAGGGCCCCGGGAATCTCACTGGACTTG	1260
Db		1550	AGTGCACCCCTGGTGATGATACAGACCCCATTCAAAAGGGCCCCGGGAATCTCACTGGACTTG	1609
OY		1261	GAATGATCACTGTCCACACATCTCATCACTTACAGAGCAGACCTCGTGATCCGACTGGCG	1320
Db		1610	GAGATGATCACCGCTCAACACATCTCATCACTTACAGAGCAGACCTCGTGATCCGACTGGCG	1669
OY		1321	ATATATGTGTGTGCAATACCAATTC	1344
Db		1670	ATATACGTGTCCAGATATCCGCTC	1693
RESULT 13				
LOCUS	AF137350	2304 bp	mRNA	linear ROD 18-JAN-2000
DEFINITION	Rattus norvegicus embryonic vascular EGF repeat-containing protein			
ACCESSION	AF137350			
VERSION	AF137350.1	GI:4583508		
KEYWORDS	.			
SOURCE	Rattus norvegicus.			
ORGANISM	Rattus norvegicus			
	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 2304)			
AUTHORS	Kowal,R.C., Richardson,J.A., Miano,J.M. and Olson,E.N.			
TITLE	EVPC, a novel epidermal growth factor-like repeat-containing			
JOURNAL	protein upregulated in embryonic and diseased adult vasculature			
PUBLISHED	Circ. Res. 84 (10), 1166-1176 (1999)			
REVIEWED	99278197			
AUTHORS	10347091			
	2 (bases 1 to 2304)			
	Kowal,R.C., Richardson,J.A., Miano,J.M. and Olson,E.N.			
TITLE	Direct Submission			
	Submitted (25-MAR-1999) Departments of Molecular Biology and			

FEATURES	source	CDs
Oncology, Internal Medicine and Pathology, UT, Southwestern Medical Center, Dallas, 5323 Harry Hines Blvd., Dallas, TX 75235-9148	Location/Qualifiers	
1. .2304	/organism="Rattus norvegicus"	
/db.xref="taxon:10116"		
153. .1499	/note="contains calcium-binding Egr repeats"	
/codon_start=1		
/product="embryonic vascular EGF repeat-containing protein EVEC"		
/protein_id="AAD25101.1"		
/db.xref="GI:4583509"		
/translation="MPGLRIITVITLALMLPHGGMNOOCTNGEDLDROGCOGLDILH ECRITPEACRGMNVONGCYLCIPRNPYRKRSPYSPYSTSGPPLAPAPVPMAS NYRPLISPLVRFECQOMDEBNQCVDDVDCADSHQCNFTQICINTEGYTSCDGVW LLEGGCLDIDCRVGYCOQLCANVPYSCVGNHFTLNDGRCCODVNEETENPCV QYVNTVYSFICRCDPGYEPEDIDHCSMDGSGSEFLCMLDNGEVPYSGFSCCPGE YVLVDNRSCODINCEHNRHTCTPLQCYMLDGGFGLIDYVEEPEYLLIGDNRMKO PAETNGKDPFFTLIFRMDVYSGRSVADIFQMATRYRGAYVIFQIKSGNGREY YMRQTGPSATLVMPRIKPGPRDLDLEMITVMTVINFRSSVIRILRITYSOLP"		
BASE COUNT	539 a 620 c 550 g 595 t	
ORIGIN		
Query Match	82.0%; Score 1102.4; DB 10; Length 2304;	
Best Local Similarity	88.8%; Pred. No. 1.5e-308;	
Matches 1193; Conservative	0; Mismatches 151; Indels 0; Gaps 0;	
QY	1 ATGCCAGAAATAAAAGGATACCTGACAGTTCACATTCCTGGCTCTCTGCTTCCAAAGCCT	60
DB	153 ATGCCAGATTTAAAAGATATCTACCTGTCACCATCTCGGCACTGTGGCTTCCACATCT	212
QY	61 GGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGATTCGCCAATCGACAGACGT	120
DB	213 GGGATGCACAGCAACAGTGCACCAAGGGCTTTGACCTGACCTGCAGCCACAGACAGAGT	272
QY	121 TTGATATATGATGATGGCAACCATCCCGAGGCGCTCCGAGAGACATGATGTGT	180
DB	273 TTGATATATGATGATGGCAACCATCCCGAGGCGCTCCGAGAGACATGATGTGT	332
QY	181 AACCAAAATGGCGGGATTTTATGCAATCCCGGAGCAAAACCTGTGATCGAGGCGCTTAC	240
DB	333 AACCAAGATGGCGGGATTTGATGATCCCTCGAACAACCGATGATTCAGAGGCGCTTAC	392
QY	241 TCGAACCCTTACTGAGACCCCTACTCAAGTCCGATCCAGACAGCTGCCCAACCACTCA	300
DB	393 TCGAATCCCTTACTGATCTACTCAAGGCGCAATCCAGACAGCGCCGACACCACTGTC	452
QY	301 GCTCAAACTATCCAGATCTCCAGGCGCTTTATGCGCGCTTGGATATCCAGATGAT	360
DB	453 GCTTCCAACTATCCAGATCTCCAGGCGCTTTGATGCGCTTGGATATCCAGATGAT	512
QY	361 GAAAGCAACCAATGTGTGATGTGACAGATGTGCAACAGATTTCCACACAGTCMACCC	420
DB	513 GAAGGCAACCAAGTGTGTGATGTGACAGATGTGCGACAGATTTACACACAGTGAACCT	572
QY	421 ACCAGATTCGATCAATACTGAAAGGGGGGTATACACCGTCCTCCAGACAGATATGG	480
DB	573 ACCAGATTCGATCAACAGGAAAGGGGTATACCGTCCTCCAGATGATGGGTATGG	632
QY	481 CTCTGGAAGGCAAGTCTTAGACATTTGAATGTGCTATGATGTTACTCCAGAGCTC	540
DB	633 CTCTGGAAGGCAAGTCTTAGATATTTGAATGTGCTATGATGTTACTCCAGAGCTC	692
QY	541 TGTGCGAATTTCTCTGATCTTATCTTGTATATGCAACCTGTGTTTACCTCAATGAG	600
DB	693 TGTGCGAATTTCTCTGATCTTATCTTGTATATGCAACCTGTGTTTACCTCAATGAG	752
QY	601 GATGAAGGCTTGGCAAGATGGAAGAGATGTGCAACGAGAAACCCCTGGCGGCAAC	660
DB	753 GATGAAGGCTTGGCAAGATGGAAGAGATGTGGAACACTGAGAACCCCTGTGTTCAAGC	812
QY	661 TGCCTCAACACTTGAAGGCTTTTCAATCTGCCGCTGTGACCACAGATATGAATTTGAGAA	720


```

Db      780 GATGGCATTCACGTCAGTGTATGAGCAGATGCGCTTCCTCGAGTTCTTGTTCACAC 839
Qy      781 GAGGTGTGAACACGCGCGGACATCTTCTGCTCCTCCAGGCTACATCCTGCTG 840
Db      840 GAGTGTGTAAACGCGCGGCTCACTTCTGCTCGTCCCTCCAGGCTACGCTCTGTG 899
Qy      841 GATGACACCGAAGCGTCCAGACATCAACGAATGTGACACAGAACACACGTCGAC 900
Db      900 GATGATACCGAAGCTGCGCAGGATATCAAGAAATGAGACACGAAACACACGTCGAC 959
Qy      901 CTGACAGACAGCGCTTCAATTTACAAGGGGGCTTCAATATGATGACACCCCGCTG 960
Db      960 TCACGACAGACTTGCTACATCTACAAAGGGGGCTTCAATATGATGATCCATCAGCTGT 1019
Qy      961 GAGGAGCCCTTATCTGAGATCACTGATACACCGCTGATGTGCTCTGTGAGAACCTGGC 1020
Db      1020 GAGGAGCCCTTATCTGCTGATGTGAAACCGCGCTGATGTGCTCTGTGAGACACAGC 1079
Qy      1021 TGCAGAGACAGCGCTTCAATCTGTACCGGAGACATGAGACGCTGCTGAGACGCTCC 1080
Db      1080 TCCAGAGACAGCGCTTCAATCTGTATCGGAGATGATGTGCTGAGACGCTCC 1139
Qy      1081 GTTCCGCTGACATCTTCCAAATGACAGCCAGCCGCTACCTCGGGGCTTATTAAT 1140
Db      1140 GTTCCGCTGACATCTTCCAGATGACAGCAACACCGATACCTGCTGCTATTAAT 1199
Qy      1141 TTCAGATCAATCTGAGGATGAGGAGAGAAATTTTACATGCGGGCAACGGGCCCCATC 1200
Db      1200 TTCAGATCAATCTGAGGAGAGAGGAGGAGGCTTATGCGGGCAACGGGCCCCATC 1259
Qy      1201 AGTGCACCGCTGTGATGACAGCGCCCATCAAGGGGGCGGGAAATCCAGCTGACTG 1260
Db      1260 AGTGCACCGCTGTGATGACAGCGCCCATCAAGGGGGCGGGAAATCCAGCTGACTG 1319
Qy      1261 GAAATGATCACTGTCAACACTGTCACTCACTTCAAGAGGAGCTCCGCTGATCCGACTGG 1320
Db      1320 GAGATGATCACTGTCAACACTGTCACTCACTTCAAGAGGAGCTCCGCTGATCCGACTGG 1379
Qy      1321 ATATATGTGTCGACATCCATTC 1344
Db      1380 ATATATGTGTCGACATCCGTTT 1403

RESULT 11
LOCUS   AF112151          2478 bp      mRNA      linear      ROD 02-AUG-1999
DEFINITION Mus musculus developmental arteries and neural crest EGF-like
ACCESSION AF112151
VERSION   AF112151.1 GI:5305670
KEYWORDS
SOURCE    Mus musculus.
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE Nakamura,T., Ruiz-Lozano,P., Lindner,V., Yabe,D., Tanisaki,M.,
AUTHORS   Furukawa,Y., Kobuke,K., Tashiro,K., Lu,Z., Andon,N.U., Schaub,R.,
           Matsumori,A., Sasayama,S., Chien,K.R. and Honjo,T.
           DANCE, a novel secreted RGD protein expressed in developing,
           atherosclerotic, and balloon-injured arteries
           J. Biol. Chem. 274 (32), 22476-22483 (1999)
TITLE     JOURNAL MEDLINE PUBMED
           99357779 10428823
REFERENCE 2 (bases 1 to 2478)
AUTHORS   Nakamura,T., Yabe,D., Tashiro,K. and Honjo,T.
TITLE     Direct Submission
JOURNAL   Submitted (09-DEC-1998) Medicine, University of California San
           Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
FEATURES
           source
           1..2478
           /organism="Mus musculus"

```

```

CDS
/db_xref="taxon:10090"
320..1666
/note="DANCE; contains 6 cDEG domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="AAD1767.1"
/db_xref="GI:5305671"
/translation="MGKRLIVTILALPHRGNAGQCTNFPDIDROSGCCLDID
ECRTPEACRDMCYNQNGILICIPFTINVRKIPISNPYSTISGTPAAPPAPAS
NPFTISRLPVCRRGYMDENQGVDECDVDSHQNPICITIEGTYCTSDGV
LLEGCLDIDECRYGYCOOLCANVPGSYSCNPFGLNDGSCODVNECTENPV
QICVNTGSGFICRDPYELEEDGIIHCSDECSFSEFLQHECVDPGVSFSGCPG
VLLDNRSGODINECHNRCTSLQTCYNILQGGFCTIDISCEFFYLLIGENRCW
PAHNSCSROPFTLLYRDMVYSGRSVPADIRPQATRRYPGATYITQIKSNGREF
YKRPSPISATLMTIRPIKRPDIQDLEMTIVNTVINFQSSVIRLITVSYRPF"
BASE COUNT 604 a 656 c 575 g 643 t
ORIGIN
Query Match 82.9%; Score 1113.6; DB 10; Length 2478;
Best Local Similarity 89.3%; Pred. No 8.2e-312;
Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
Qy      1 ATCCAGGAATATAAAGATACCTACTGTTACCATTTCTGCTCTGTTCCAAAGCCT 60
Db      320 ATGCCAGATATAAAGATATAAAGATACCTACTGTTACCATTTCTGCTCTGCTCCACATCCT 379
Qy      61 GGGAAATGCACAGGACAGTGCACAGATGCGTTGACCTGATGAGCCAGTACAGACAGT 120
Db      380 GGGAAATGCACAGGACAGTGCACAGATGCGTTGACCTGATGAGCCAGTACAGACAGT 439
Qy      121 TTAGATATTGATGAATAGCCGAACCATCCCGAGAGCCCGGAGAGACATGATGTGTT 180
Db      440 CTAGATATTGATGAATAGCCGAGACATCCCTGAGAGCTTGTGTGGGACATGATGTGTC 499
Qy      181 AACCAAAATGGCGGATATTATGCAATTCGCCGAGCAAAACCTGTGATGAGAGGCGCTAC 240
Db      500 AACCAAAATGGCGGATATTATGCAATTCGCCGAGCAAAACCTGTGATGAGAGGCGCTAC 559
Qy      241 TCGAACCCCTACTCGACGCCCTTCTCAGAGTCGATGACAGACAGTGCCTCCACACTGCA 300
Db      560 TCAAAATCCCTACTCTTACATCTTCTCAGAGTCGATGACAGACAGTGCCTCCACACTGCA 619
Qy      301 GCTCAAACTATCCACAGATCTCCAGGCTCTTATATGCGCTTGTGATACCAAGATGAT 360
Db      620 GCTTCAACTATCCACAGATCTCCAGGCTCTTGTGTGCGGCTTGTGATACAGATGAT 679
Qy      361 GAAACCAACCATGTGTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGCACACCC 420
Db      680 GAAAGCAACCATGTGTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGCACACCC 739
Qy      421 ACCCGATCTGATCAATCTGAGAGCGGGTACACTGCTCGTCCGACCGAGGATATGTT 480
Db      740 ACCCGATCTGATCAATCTGAGAGAGTGTGATGATGATGATGATGATGATGATGATG 799
Qy      481 CTTTGAAGGCGCACTGTAGACATTTGATGATGATGATGATGATGATGATGATGATG 540
Db      800 CTTTGAAGGCGCACTGTAGATTTGATGATGATGATGATGATGATGATGATGATGATG 859
Qy      541 TGTGGATGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db      860 TGTGGATGTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 919
Qy      601 GATGAGAGTCTTGGCAAGATGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db      920 GATGAGAGTCTTGGCAAGATGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 979
Qy      661 TGCCTCAACACCTAGAGGCTTTTCACTCTGCGCTGAGACCAAGATATGAATGTGAGAA 720
Db      980 TGTGCAACACCTATGGCTTTTCACTCTGCGCTGAGACCAAGATATGAATGTGAGAA 1039

```

QY 1021 TGCAGAGACCGCCCTTACCATCTGTACCGGAGATGACGTGTGTACAGACGCTCC 1080
 DB 1189 TGCAGAGACCGCCCTTACCATCTGTACCGGAGATGACGTGTGTACAGACGCTCC 1248
 QY 1081 GTTCCCGCTGACATCTTCAAAATGCAAGCAGACCGCCCTGAGGCGCTATTACATT 1140
 DB 1249 GTTCCCGCTGACATCTTCAAAATGCAAGCAGACCGCCCTGAGGCGCTATTACATT 1308
 QY 1141 TTCCAGATCAAACTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGCGCCCATC 1200
 DB 1309 TTCCAGATCAAACTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGCGCCCATC 1368
 QY 1201 AGTGCACCTGTGTGATGACAGACCGCCCATCAAGGCGCGGGAATCCAGCTGACTTG 1260
 DB 1369 AGTGCACCTGTGTGATGACAGACCGCCCATCAAGGCGCGGGAATCCAGCTGACTTG 1428
 QY 1261 GAAATGATCACTGTCAACACTGTCACTCAAGAGGACGCTCGGTATCCGACTGCGG 1320
 DB 1429 GAAATGATCACTGTCAACACTGTCACTCAAGAGGACGCTCGGTATCCGACTGCGG 1488
 QY 1321 ATATATGTGTGAGTACCATTC 1344
 DB 1489 ATATATGTGTGAGTACCATTC 1512

RESULT 10
 BC006636 2230 bp mRNA linear ROD 07-ANG-2002
 LOCUS Mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,
 complete cds.
 ACCESSION BC006636
 VERSION BC006636.1 GI:13879321
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Submitted (27-APR-2001) National Institutes of Health, Mammalian
 TITLE Gene Collection (MGC), Cancer Genomics Office, National Cancer
 JOURNAL Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 REMARK
 COMMENT Contact: MGC help desk
 Email: cgap@ncl.nih.gov
 Tissue Procurement: Lohar Hennighausen Ph.D., Robin Humphreys
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

FEATURES
 source
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 6 Row: f Column: 8
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6753823.
 Location/Qualifiers

1. 2230
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="C57BL/6J"
 /clone="MGC:5656 IMAGE:3482574"
 /tissue="MGC:Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."

CDS
 /clone_11b="NCL_GAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 60..1406
 /codon_start=1
 /product="Fibulin 5"
 /protein_id="AAH06636.1"
 /db_xref="GI:13879322"
 /db_xref="locusID:23876"

BASE COUNT 562 a 578 c 505 g 585 t
 ORIGIN

Query Match 82.9% Score 1113.6; DB 10; Length 2230;
 Best Local Similarity 89.3%; Pred No. 8.2e-312;
 Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 1 ATGCCAGGAATAAAGAGTACTACTGTACCATCTGTCTCTCTCAACCCCT 60
 DB 60 ATGCCAGGAATAAAGAGTACTACTGTACCATCTGTCTCTCTCAACCCCT 119
 QY 61 GGGATGACACAGGACATGACACGAATGCTTTGACCTGATGCGGACGACAGCTGT 120
 DB 120 GGGATGACACAGGACATGACACGAATGCTTTGACCTGATGCGGACGACAGCTGT 179
 QY 121 TTGATATTTGATGATGCGGACCAACATCCCGGCGCTGCGGAGAGACATGATGTGT 180
 DB 180 CTGATATTTGATGATGCGGACCAACATCCCGGCGCTGCGGAGAGACATGATGTGT 239
 QY 181 AACCAAAATGGCGGGTATTTATGATCCCGGACAAACCTGTGTATGAGGGCCCTAC 240
 DB 240 AACCAAAATGGCGGGTATTTATGATCCCGGACCAACCTGTGTATGAGGGCCCTAC 299
 QY 241 TCGAACCCTTCTGACCCCTTCTGACAGTCCGTACCCAGACAGCTCCCGACACTCTCA 300
 DB 300 TCGAACCCTTCTGACCCCTTCTGACAGTCCGTACCCAGACAGCTCCCGACACTCTCA 359
 QY 301 GCTCAAACTATCCAGCATCTCCAGGCTCTTATATGCGGCTTGTGATACCAAGATGAT 360
 DB 360 GCTCAAACTATCCAGCATCTCCAGGCTCTTATATGCGGCTTGTGATACCAAGATGAT 419
 QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCACAGAGTCAACCCC 420
 DB 420 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCACAGAGTCAACCCC 479
 QY 421 ACCAGATCTGATCAATACAGAGGCGGTACCTGCTCTGACCCAGCATATGTG 480
 DB 480 ACCAGATCTGATCAATACAGAGGCGGTACCTGCTCTGACCCAGCATATGTG 539
 QY 481 CTTTGTGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 540 CTTTGTGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
 QY 541 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 600 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
 QY 601 GATGGAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 660 GATGGAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
 QY 661 TGTGCAACCACTGAGCTCTTTCATCTGCGGCTGTGACCCAGATATTAACCTGAGGAA 720
 DB 720 TGTGCAACCACTGAGCTCTTTCATCTGCGGCTGTGACCCAGATATTAACCTGAGGAA 779
 QY 721 GATGCGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 901 CTGCACAGACGCTGCTACAAATTTACAAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGT 960
|||||
Db 1350 CTGCACAGACGCTGCTACAAATTTACAAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGT 1409
QY 961 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCCCTGGC 1020
|||||
Db 1410 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTCTGCTGAGAACCCCTGGC 1469
QY 1021 TGCAGAGACACGCGCTTACCATCTTGTACCGGAGATGAGCGTGTGTGAGAGAGCTTC 1080
|||||
Db 1470 TGCAGAGACACGCGCTTACCATCTTGTACCGGAGATGAGCGTGTGTGAGAGAGCTTC 1529
QY 1081 GTTCCGCTGACATCTTCAAAATGCAAGCAGACCGCTACCCCTGAGGCTTATACAT 1140
|||||
Db 1530 GTTCCGCTGACATCTTCAAAATGCAAGCAGACCGCTACCCCTGAGGCTTATACAT 1589
QY 1141 TTCCAAATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCATC 1200
|||||
Db 1590 TTCCAAATCAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCATC 1649
QY 1201 AGTGCACCTGGTGTATGACACGCGCCATCAAAAGGGGCCGGAATTCAGCTGTGACATTG 1260
|||||
Db 1650 AGTGCACCTGGTGTATGACACGCGCCATCAAAAGGGGCCGGAATTCAGCTGTGACATTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTTCATCACTTTCAGAGGACAGCTCCGTGATCGAGCTGG 1320
|||||
Db 1710 GAAATGATCACTGTCAACACTGTTCATCACTTTCAGAGGACAGCTCCGTGATCGAGCTGG 1769
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
|||||
Db 1770 ATATATGTGTGCGAGTACCATTC 1793

RESULT 9
AF093118 2019 bp mRNA 1linear PRI 01-OCT-1998
LOCUS Homo sapiens UP50 mRNA, complete cds.
DEFINITION AF093118
ACCESSION AF093118.1 GI:3676821
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Zemel, R., Sholto, O. and Shaul, Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann
Institute of Science, Rehovot 76100, Israel
FEATURES
source
1. 2019
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="urine"
1. 168
159..1515
/note="50KD: similar to human SI-5 encoded by GenBank
Accession Number U03877; possible membrane and
extracellular matrix protein; urine p50 protein"
/codon_start=1
/product="UP50"
/protein_id="AAC62107.1"
/db_xref="GI:3676822"

5'UTR
CDS
3'UTR
BASE COUNT 478 a 585 c 490 g 465 t 1 others

ORIGIN
Query Match 98.8% Score 1328: DB 9: Length 2019;
Best Local Similarity 99.3%: Pred. No. 0; Mismatches 10; Indels 0; Gaps 0;
Matches 1334; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ATCCAGAGATTAAGAGATCTACTGTTACATCTTGGCTCTGCTTCCAAAGCCCT 60
Db 169 ATCCAGAGATTAAGAGATCTACTGTTACATCTTGGCTCTGCTTCCAAAGCCCT 228
QY 61 GGGATGACACAGGACAGTGCAGAAATGCTTGTGACCTGTGATGCCAGTACAGAGTCT 120
Db 229 GGGATGACACAGGACAGTGCAGAAATGCTTGTGACCTGTGATGCCAGTACAGAGTCT 288
QY 121 TTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 289 TTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 348
QY 181 AACCAAAATGCGGGATTTTATGATTCACCTCCGCAAAACCTGTATGAGAGGCTTAC 240
Db 349 AACCAAAATGCGGGATTTTATGATTCACCTCCGCAAAACCTGTATGAGAGGCTTAC 408
QY 241 TCGAACCCTACTGACCCCTACTCAGTCCGTATCCAGAGAGTCCCTCCACACTCTCA 300
Db 409 TCGAACCCTACTGACCCCTACTCAGTCCGTATCCAGAGAGTCCCTCCACACTCTCA 468
QY 301 GCTCCAAATCTCCAGATCTCCAGGCTCTTATATGCGCTTGGATTAACAGATGAT 360
Db 469 GCTCCAAATCTCCAGATCTCCAGGCTCTTATATGCGCTTGGATTAACAGATGAT 528
QY 361 GAAAGCAACATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCACACCC 420
Db 529 GAAAGCAACATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCACACCC 588
QY 421 ACCCAGATCTGCATCAATTAAGAGGCGGTACACCTGCTCTGACAGGAGGATTTGG 480
Db 589 ACCCAGATCTGCATCAATTAAGAGGCGGTACACCTGCTCTGACAGGAGGATTTGG 648
QY 481 CTTCTGGAAGGCGAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 649 CTTCTGGAAGGCGAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 708
QY 541 TGTGCAATGTTCTGTGATCTTATCTTGTGATGATGATGATGATGATGATGATGATGAT 600
Db 709 TGTGCAATGTTCTGTGATCTTATCTTGTGATGATGATGATGATGATGATGATGATGAT 768
QY 601 GATGGAAGTCTTGTGCAAGATGTGAAGAGTGTGCGACAGCAACCCCTGCTGCAACC 660
Db 769 GATGGAAGTCTTGTGCAAGATGTGAAGAGTGTGCGACAGCAACCCCTGCTGCAACC 828
QY 661 TGCCTCAACACCTACGCTCTTATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 829 TGCCTCAACACCTACGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 888
QY 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 889 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
QY 781 GAGTGTGTGACACAGCCCGGACATCTTGTCTGCTGCTCCAGGCTTACATCTGCTG 840
Db 949 GAGTGTGTGACACAGCCCGGACATCTTGTCTGCTGCTCCAGGCTTACATCTGCTG 1008
QY 841 GATGACAACGGAAGCTCCCAAGCATACGAAGTGTGACAGGAACACACAGTGCAC 900
Db 1009 GATGACAACGGAAGCTCCCAAGCATACGAAGTGTGACAGGAACACACAGTGCAC 1068
QY 901 CTGCACAGACGCTGCTCAATTTCAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGT 960
Db 1069 CTGCACAGACGCTGCTCAATTTCAAGGGGGCTTCAAAATGATCAGACCCCATCCGCTGT 1128
QY 961 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGGC 1020
Db 1129 GAGGACCTTATCTGAGGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGGC 1188

QY 601 GATGGAAGGCTCTGCCAAGATGTGAACGAGTGTGCCAACCAGAAACCCCTGGCTGCAACC 660
DB 1050 GATGGAAGGCTCTGCCAAGATGTGAACGAGTGTGCCAACCAGAAACCCCTGGCTGCAACC 1109
QY 661 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGAA 1169
QY 721 GATGGCGTTCATTGCAAGTATGAGAGAGTGCAGCTTCTGTGAGTCTCTGCCAAT 780
DB 1170 GATGGCGTTCATTGCAAGTATGAGAGAGTGCAGCTTCTGTGAGTCTCTGCCAAT 1229
QY 781 GAGTGTGTAACACGAGCCGCGACATATCTTGCCTCGCCCTCCAGAGGTACATCTGCG 840
DB 1230 GAGTGTGTAACACGAGCCGCGACATATCTTGCCTCGCCCTCCAGAGGTACATCTGCG 1289
QY 841 GATGACACCGAAGCTGCCAAGATCAGCAAGATGTGAGCAGAGAACACACGTCGAC 900
DB 1290 GATGACACCGAAGCTGCCAAGATCAGCAAGATGTGAGCAGAGAACACACGTCGAC 1349
QY 901 CTGCAACACACGCTGCTACATTTTCAAGGGGCTTCAATGATCGACCCATCGCTGT 960
DB 1350 CTGCAACACACGCTGCTACATTTTCAAGGGGCTTCAATGATCGACCCATCGCTGT 1409
QY 961 GAGAGACCTTATCTGAGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGAC 1020
DB 1410 GAGAGACCTTATCTGAGATCAGTATACCGCTGTATGTCTGCTGAGAACCCCTGAC 1469
QY 1021 TGCAGAGACACGAGCCCTTTTACCATCTGTGTACCGGGACATGACGTGTGTGAGAGCTTC 1080
DB 1470 TGCAGAGACACGAGCCCTTTTACCATCTGTGTGTACCGGGACATGACGTGTGTGAGAGCTTC 1529
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCGTGGGCGCTATTACAT 1140
DB 1530 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCGTGGGCGCTATTACAT 1589
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTTACATGCGGCAAAAGGCGCCATC 1200
DB 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAAATTTTACATGCGGCAAAAGGCGCCATC 1649
QY 1201 AGTGCACACCTGCTGATGACACGCGCCCATCAAAAGGGCGCGGAAATCAGCTGACCTTG 1260
DB 1650 AGTGCACACCTGCTGATGACACGCGCCCATCAAAAGGGCGCGGAAATCAGCTGACCTTG 1709
QY 1261 GAAATGATCACTGTCAACACTGTATCATCACTTGTGAGGAGAGCTCGGTATCGAGTGGG 1320
DB 1710 GAAATGATCACTGTCAACACTGTATCATCACTTGTGAGGAGAGCTCGGTATCGAGTGGG 1769
QY 1321 ATATATGTGTGAGTACCATTC 1344
DB 1770 ATATATGTGTGAGTACCATTC 1793
RESULT 8
AX403659 2609 bp DNA linear PAT 14-JUN-2002
LOCUS AX403659
DEFINITION Sequence 14 from Patent W00077037.
ACCESSION AX403659
VERSION AX403659.1 GI:21437116
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Ashkenazi, A., Baker, K., Botstein, D., Desnoyers, L., Eaton, D.L.,
Ferrara, N., Fong, S., Gao, W.Q., Gerber, H., Gertlisen, M.E.,
Goddard, A., Godowski, P., Gurney, A., Kijavlin, I.J., Mather, J.,
Napier, M., Pan, J., Paoni, N., Roy, M., Tumas, D., Watanabe, C.,
Williams, P.M., Wood, W.I. and Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same

JOURNAL Patent: WO 0077037-A 14 21-DEC-2000;
Genentech Inc. (US)
FEATURES Location/Qualifiers
source 1..2609 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 622 a 718 c 622 g 647 t
ORIGIN
Query Match 99.98; Score 1342.4; DB 6; Length 2609;
Best Local Similarly 99.98; Pred. No. 0;
Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCAGAAATTAAGATACATCAGTGTCCATTTCTGCGCTCTGCTTCCAAAGCCCT 60
DB 450 ATGCCAGAAATTAAGATACATCAGTGTCCATTTCTGCGCTCTGCTTCCAAAGCCCT 509
QY 61 GGAATGACACAGGACAGTGCAGCAATGGCTTTGACCTGATCCGCACTCAGACAGTGT 120
DB 510 GGAATGACACAGGACAGTGCAGCAATGGCTTTGACCTGATCCGCACTCAGACAGTGT 569
QY 121 TTGATATTTATGATATCCGCAACCATCCCGAGGCTTCCGAGAGACATGATGTGT 180
DB 570 TTGATATTTATGATATCCGCAACCATCCCGAGGCTTCCGAGAGACATGATGTGT 629
QY 181 AACCAAAATGGGGATTTTATGATTCGCCGAAACCCCTGTATGAGAGGCGCTTAC 240
DB 630 AACCAAAATGGGGATTTTATGATTCGCCGAAACCCCTGTATGAGAGGCGCTTAC 689
QY 241 TCGAACCCTTACTGACCCCTTACTCAGAGTCCGTACCAAGACAGTGTGCCACACTTCA 300
DB 690 TCGAACCCTTACTGACCCCTTACTCAGAGTCCGTACCAAGACAGTGTGCCACACTTCA 749
QY 301 GCTCAAAATATCCAGATCTCCAGGCTCTTATATGCGGCTTGTGATACAGATGAT 360
DB 750 GCTCAAAATATCCAGATCTCCAGGCTCTTATATGCGGCTTGTGATACAGATGAT 809
QY 361 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAAGATTTCCACAGTGCACACCC 420
DB 810 GAAAGCAACCAATGTGTGATGTGAGAGTGTGCAAGATTTCCACAGTGCACACCC 869
QY 421 ACCCAATCTGATCAATACTGAAGCGGGGTACACCTCTCTGCAAGCAGATTTGG 480
DB 870 ACCCAATCTGATCAATACTGAAGCGGGGTACACCTCTCTGCAAGCAGATTTGG 929
QY 481 CTTCTGGAAGGCAAGCTTGTAGATGATGAATGTGCTATGTTACTGCCAGAGCTC 540
DB 930 CTTCTGGAAGGCAAGCTTGTAGATGATGAATGTGCTATGTTACTGCCAGAGCTC 989
QY 541 TGTGCAATGTTCCTGATCCTATTTGTACATGCAACCCGTGTTTACCTCAATGAG 600
DB 990 TGTGCAATGTTCCTGATCCTATTTGTACATGCAACCCGTGTTTACCTCAATGAG 1049
QY 601 GATGGAAGGCTCTGCCAAGATGTGAACGAGTGTGCCAACCAGAAACCCCTGGCTGCAACC 660
DB 1050 GATGGAAGGCTCTGCCAAGATGTGAACGAGTGTGCCAACCAGAAACCCCTGGCTGCAACC 1109
QY 661 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGAA 720
DB 1110 TGCCTGCAACACCTACGAGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAGAA 1169
QY 721 GATGGCGTTCATTGCAAGTATGAGAGAGTGCAGCTTCTGTGAGTCTCTGCCAAT 780
DB 1170 GATGGCGTTCATTGCAAGTATGAGAGAGTGCAGCTTCTGTGAGTCTCTGCCAAT 1229
QY 781 GAGTGTGTAACACGAGCCGCGACATATCTTGCCTCGCCCTCCAGAGGTACATCTGCG 840
DB 1230 GAGTGTGTAACACGAGCCGCGACATATCTTGCCTCGCCCTCCAGAGGTACATCTGCG 1289
QY 841 GATGACACCGAAGCTGCCAAGATCAGCAAGATGTGAGCAGAGAACACACGTCGAC 900
DB 1290 GATGACACCGAAGCTGCCAAGATCAGCAAGATGTGAGCAGAGAACACACGTCGAC 1349

OY	241	TCGAAACCCCTAC	TGACGACCCCTACT	CAAGGTCCTTAC	CCAGACACTG	CCCACTCTCA	300
Db	622	TCGAAACCCCTACT	CGACCCCTACT	CAAGGTCCTTAC	CCAGACACTG	CCCACTCTCA	681
OY	301	GCTCCAAACAT	TATCCACGAT	CTCCAGGCTCTT	ATATGCGCGCTT	TGATACCGAGAT	360
Db	682	GCTCCAAACAT	TATCCACGAT	CTCCAGGCTCTT	ATATGCGCGCTT	TGATACCGAGAT	741
OY	361	GAAGGCAACCA	TATGTGTGGAT	GTGGACAGAT	GTGCAACAGAT	TCCACAGTGCAACCC	420
Db	742	GAAGGCAACCA	TATGTGTGGAT	GTGGACAGAT	GTGCAACAGAT	TCCACAGTGCAACCC	801
OY	421	ACCCAGATCTG	CAATACTGAAG	CGGGTACACCTG	CTCGACCGACG	AGATATTTGG	480
Db	802	ACCCAGATCTG	CAATACTGAAG	CGGGTACACCTG	CTCGACCGACG	AGATATTTGG	861
OY	481	CTTCTGGAAG	CGCCAGTGTG	ACATTTGAT	TGAATGTGG	TATGTACTGACGACGCTC	540
Db	862	CTTCTGGAAG	CGCCAGTGTG	ACATTTGAT	TGAATGTGG	TATGTACTGACGACGCTC	921
OY	541	TGAGCAATGT	TCCTGGATCT	TATTCCTT	TACATGTAAC	CCCTGGTTTAC	600
Db	922	TGAGCAATGT	TCCTGGATCT	TATTCCTT	TACATGTAAC	CCCTGGTTTAC	981
OY	601	GATGGAAGTCT	TGGCCAMAT	GTGAACGAG	TGTGCACCGA	AGAACCCCTGCTGCAAAAC	660
Db	982	GATGGAAGTCT	TGGCCAMAT	GTGAACGAG	TGTGCACCGA	AGAACCCCTGCTGCAAAAC	104
OY	661	TGCGTCAAC	ACCTACGCGCT	CTTTCATCT	CTGCGCTGT	GAACCCAGATATG	720
Db	1042	TGCGTCAAC	ACCTACGCGCT	CTTTCATCT	CTGCGCTGT	GAACCCAGATATG	1101
OY	721	GATGGCGTT	CATTGCAAG	TATATGAG	ACGAGTAC	GTCTGTCGATG	780
Db	1102	GATGGCGTT	CATTGCAAG	TATATGAG	ACGAGTAC	GTCTGTCGATG	1161
OY	781	GAGTGTGTA	CAACGACCG	CGGACAT	CTTCTGCTT	GGCCTCGACAGGCTAC	840
Db	1162	GAGTGTGTA	CAACGACCG	CGGACAT	CTTCTGCTT	GGCCTCGACAGGCTAC	1221
OY	841	GATGCAAC	CCGAAAGCT	TGCGCAAG	ACATCAAG	AATGTGAGACAGGAAC	900
Db	1222	GATGCAAC	CCGAAAGCT	TGCGCAAG	ACATCAAG	AATGTGAGACAGGAAC	1281
OY	901	CTGACGAC	AGACGTGCT	TACAAATTT	TACAAAGGGGGCTT	CAATGATC	960
Db	1282	CTGACGAC	AGACGTGCT	TACAAATTT	TACAAAGGGGGCTT	CAATGATC	1341
OY	961	GAGGAGCCT	TATCTAG	GATGATG	ATAAACCGCT	TATGTCTGCTG	1020
Db	1342	GAGGAGCCT	TATCTAG	GATGATG	ATAAACCGCT	TATGTCTGCTG	1401
OY	1021	TGCAAGAC	CAACGCTTT	TACCATCT	TGTATAC	GGGACATG	1080
Db	1402	TGCAAGAC	CAACGCTTT	TACCATCT	TGTATAC	GGGACATG	1461
OY	1081	GTTCCCGG	TGACATCTT	TCCAAATG	CAAGCCAC	CGCTTAC	1140
Db	1462	GTTCCCGG	TGACATCTT	TCCAAATG	CAAGCCAC	CGCTTAC	1521
OY	1141	TTCCAGAT	CAAAATCT	TGGGAAT	GAGGGGAG	AGAAATTTT	1200
Db	1522	TTCCAGAT	CAAAATCT	TGGGAAT	GAGGGGAG	AGAAATTTT	1581
OY	1201	AGTCCCA	CCCTGGT	GATGAC	ACGCCCCAT	CAAGGGGCC	1260
Db	1582	AGTCCCA	CCCTGGT	GATGAC	ACGCCCCAT	CAAGGGGCC	1641
OY	1261	GAATGAT	CACTGTC	CAACACT	GTGATCA	CACTGAGGCA	1320
Db	1642	GAATGAT	CACTGTC	CAACACT	GTGATCA	CACTGAGGCA	1701
OY	1321	ATATATGT	GTGCGAG	TACCACTT	C	1344	

Db	1702	AFATATGTGTGCAGTACCACTTC	1725
RESULT 7			
LOCUS	AX376340	2609 bp	DNA
DEFINITION	Sequence 407 from Patent WO0168848.		linear
ACCESSION	AX376340		PAT 01-MAR-2002
VERSION	AX376340.1	GI:19170561	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Baker, K.P., Chen, J., Desnoyers, L., Goddard, A., Godowsky, P.J., Zhang, Z., Gurney, A.L., Pan, J., Smith, V., Watanabe, C.K., Wood, W.I. and the same		
TITLE	Secreted and transmembrane polypeptides and nucleic acids encoding the same		
JOURNAL	Patent: WO 0168848-A 407 20-SEP-2001; Genentech, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..2609		
BASE COUNT	622 a 718 c 622 g 647 t		
ORIGIN			
Query Match	99.98%	Score 1342.4	DB 6; Length 2609;
Best Local Similarity	99.98%	Pred. No. 0;	
Matches 1343;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	ATGCCAGGATATATAAGATACACTGCTTACCATTCCTGGCTCTGCTTCCAAAGCCCT	60
Db	450	ATGCCAGGATATATAAGATACACTGCTTACCATTCCTGGCTCTGCTTCCAAAGCCCT	509
QY	61	GGGAATGCACAGCAGCAGTGCAGCAATGGCTTTGACCTGGATGGCAGTACAGTGT	120
Db	510	GGGAATGCACAGCAGCAGTGCAGCAATGGCTTTGACCTGGATGGCAGTACAGTGT	569
QY	121	TTAGATATTGATGAATATCCGACACATCCCGAGGCTCCGAGAGACATGATGTCTGT	180
Db	570	TTAGATATTGATGAATATCCGACACATCCCGAGGCTCCGAGAGACATGATGTCTGT	629
QY	181	AACCAAAATGGCGGATTTATGATCTCCCGACAAACCTGTGTATGAGAGGCCCTAC	240
Db	630	AACCAAAATGGCGGATTTATGATCTCCCGACAAACCTGTGTATGAGAGGCCCTAC	689
QY	241	TGGAACCCCTACTCGACGCCCTCTACTGAGTCCGATACCCAGACGTCGCCCACTCTCA	300
Db	690	TGGAACCCCTACTCGACGCCCTCTACTGAGTCCGATACCCAGACGTCGCCCACTCTCA	749
QY	301	GCTCCAAATATCCACAGATCTCCAGGCTCTTATATGCCCTTTGATATACAGATGAT	360
Db	750	GCTCCAAATATCCACAGATCTCCAGGCTCTTATATGCCCTTTGATATACAGATGAT	809
QY	361	GAAGAAGAACCAATGTGTGATGTGAGACGATGTGCAACAGATTTCCACAGTGCACCC	420
Db	810	GAAGAAGAACCAATGTGTGATGTGAGACGATGTGCAACAGATTTCCACAGTGCACCC	869
QY	421	ACCAGATCTGCATCAATACTGAAAGCGGTGTACACCTGCTCCACAGAGGATTTGG	480
Db	870	ACCAGATCTGCATCAATACTGAAAGCGGTGTACACCTGCTCCACAGAGGATTTGG	929
QY	481	CTTCTGGAAGGCCAGTGTCTTACAGATGATGAATGTCGTATGTTACTGACAGCTC	540
Db	930	CTTCTGGAAGGCCAGTGTCTTACAGATGATGAATGTCGTATGTTACTGACAGCTC	989
QY	541	TGTGGGATGTCTCCAGATCTTATCTTGTGATATGCAACCGCTGTTTACCTCAATAG	600
Db	990	TGTGGGATGTCTCCAGATCTTATCTTGTGATATGCAACCGCTGTTTACCTCAATAG	1049

AUTHORS Bandman, O., Corley, N.C. and Guegler, K.J.
 TITLE Human extracellular matrix proteins
 JOURNAL Patent: US 5872234-A 2 16-FEB-1999;
 FEATURES Location/Qualifiers
 Source 1..2550
 /organism="unknown"

BASE COUNT 623 a 690 c 596 g 641 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATGCCAGATATATAAGATACATCTACATGTCATCTGCTCTCTGCTTCCAGCCCT 60
382 ATGCCAGATATATAAGATACATCTACATGTCATCTGCTCTCTGCTTCCAGCCCT 441
61 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTGACCTGATCGCCAGTCAAGACAGT 120
442 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTGACCTGATCGCCAGTCAAGACAGT 501
121 TTAGATATATGATGAATATGCGAATACCTCCCGAGGCTTCGAGAGACATGATGTGT 180
502 TTAGATATATGATGAATATGCGAATACCTCCCGAGGCTTCGAGAGACATGATGTGT 561
181 AACCAAAATGGCGGATATTTATGATCTCCCGAACAACCTGTGTATCGAGGCGCTAC 240
562 AACCAAAATGGCGGATATTTATGATCTCCCGAACAACCTGTGTATCGAGGCGCTAC 621
241 TCGAACCCCTACTGCAGCCCTCTACTCAGTCCGATACCCAGAGCTGCCCAACACTCT 300
622 TCGAACCCCTACTGCAGCCCTCTACTCAGTCCGATACCCAGAGCTGCCCAACACTCT 681
301 GCTCCAACTATCCAGATCTCAGGCGCTTATATGCGCTTGGATATACAGATGAT 360
682 GCTCCAACTATCCAGATCTCAGGCGCTTATATGCGCTTGGATATACAGATGAT 741
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACACAGTGCNA 420
742 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACACAGTGCNA 801
421 ACCCAGATGTGCATCACTACTGAAGGGGGTACACCTGCTCTCTGACCCAGAGATTT 480
802 ACCCAGATGTGCATCACTACTGAAGGGGGTACACCTGCTCTCTGACCCAGAGATTT 861
481 CTCTCTGGAAGGCGAGTGTGATGATGAATGTGCTATGCTTACTGCGACAGGCTC 540
862 CTCTCTGGAAGGCGAGTGTGATGATGAATGTGCTATGCTTACTGCGACAGGCTC 921
541 TGTGCAATGTCTCTGATCTTATCTTGTATGATGCAACCTGTGTTTACCCTCAATG 600
922 TGTGCAATGTCTCTGATCTTATCTTGTATGATGCAACCTGTGTTTACCCTCAATG 981
601 GATGGAAGTCTTTCGCAAGATGTGAGAGAGTGTGCGACCGGAACCCCTGGGTGCA 660
982 GATGGAAGTCTTTCGCAAGATGTGAGAGAGTGTGCGACCGGAACCCCTGGGTGCA 1041
661 TCGCTCAACACCTACGCGCTTTCATCTGCGCTGTACCCAGAGATATGAATTGAG 720
1042 TCGCTCAACACCTACGCGCTTTCATCTGCGCTGTACCCAGAGATATGAATTGAG 1101
721 GATGCGCTTCATTTGAGTATGAGAGAGTGTGCGCTTTCATCTGCGCTTCGCAACAT 780
1102 GATGCGCTTCATTTGAGTATGAGAGAGTGTGCGCTTTCATCTGCGCTTCGCAACAT 1161
781 GATGCTGTAACACAGGCGCGACATATCTTCTGCTCCGCTCCAGAGTATATCTGTG 840
1162 GATGCTGTAACACAGGCGCGACATATCTTCTGCTCCGCTCCAGAGTATATCTGTG 1221
841 GATGACAAACGAGAGTGTGCAAGAGATCAAGAGATGTAGACAGAGAACACACAGTGC 900
1222 GATGACAAACGAGAGTGTGCAAGAGATCAAGAGATGTAGACAGAGAACACACAGTGC 1281

```

```

QY 901 CTGACAGACAGTGTCTCAATTTTACAGGGGGCTTCAATGATGACGCCATCCGCTG 960
DB 1282 CTGACAGACAGTGTCTCAATTTTACAGGGGGCTTCAATGATGACGCCATCCGCTG 1341
QY 961 GAGAGGCTTATCTGAGAGTCAAGTAAACCGCTGTATGTGTCTGTGAGAACCTGTGC 1020
DB 1342 GAGAGGCTTATCTGAGAGTCAAGTAAACCGCTGTATGTGTCTGTGAGAACCTGTGC 1401
QY 1021 TGCAGAGACAGGCCCTTACCATTCTTGTACCGGGACATGAGAGTGTGTGACAGAGCTGC 1080
DB 1402 TGCAGAGACAGGCCCTTACCATTCTTGTACCGGGACATGAGAGTGTGTGACAGAGCTGC 1461
QY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCCCTGAGGCGCTATTACAT 1140
DB 1462 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCCCTGAGGCGCTATTACAT 1521
QY 1141 TTCCAGATCAATCTGGAATGAGGGACAGAAATTTTACATGCGGCAACGGGCCCATC 1200
DB 1522 TTCCAGATCAATCTGGAATGAGGGACAGAAATTTTACATGCGGCAACGGGCCCATC 1581
QY 1201 AGTCCACCCCTGTGATGATGACAGGCCCATCAAGGGCGCCGGGAATTCAGCTGACCTTG 1260
DB 1582 AGTCCACCCCTGTGATGATGACAGGCCCATCAAGGGCGCCGGGAATTCAGCTGACCTTG 1641
QY 1261 GAAATGATCACTGTCAACACTGTGATCAACTTTCAGAGGACAGCTCCGTGATCCGACTGCG 1320
DB 1642 GAAATGATCACTGTCAACACTGTGATCAACTTTCAGAGGACAGCTCCGTGATCCGACTGCG 1701
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1702 ATATATGTGTGCGAGTACCATTC 1725

```

RESULT 6
 AR173204 2550 bp DNA linear PAT 17-DEC-2001
 LOCUS AR173204
 DEFINITION Sequence 2 from patent US 6303765.
 ACCESSION AR173204
 VERSION AR173204.1 GI:17912695
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2550)
 AUTHORS Bandman, O., Corley, N.C. and Guegler, K.J.
 TITLE Human extracellular matrix proteins
 JOURNAL Patent: US 6303765-A 2 16-OCT-2001;
 FEATURES Location/Qualifiers
 Source 1..2550
 /organism="unknown"

BASE COUNT 623 a 690 c 596 g 641 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATGCCAGATATATAAGATACATCTACATGTCATCTGCTCTCTGCTTCCAGCCCT 60
382 ATGCCAGATATATAAGATACATCTACATGTCATCTGCTCTCTGCTTCCAGCCCT 441
61 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTGACCTGATCGCCAGTCAAGACAGT 120
442 GGGAAATGCACAGGACAGTGCACGAAATGGCTTTGACCTGATCGCCAGTCAAGACAGT 501
121 TTAGATATATGATGAATATGCGAATACCTCCCGAGGCTTCGAGAGACATGATGTGT 180
502 TTAGATATATGATGAATATGCGAATACCTCCCGAGGCTTCGAGAGACATGATGTGT 561
181 AACCAAAATGGCGGATATTTATGATCTCCCGAACAACCTGTGTATCGAGGCGCTAC 240
562 AACCAAAATGGCGGATATTTATGATCTCCCGAACAACCTGTGTATCGAGGCGCTAC 621

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.sngc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 36 Row: n Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5453649.

FEATURES

source

1..2367
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:22412 IMAGE:4693953"
 /issue_type="Lung"
 /clone_lib="NIR_MGC_77"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 184..1530
 /codon_start=1
 /product="unknown (protein for MGC:22412)"
 /protein_id="AAH22280.1"
 /db_xref="GI:18490145"
 /translation="MGIRKRLVTLIALCLPSPGNQAOCTNCFDIDRSGOCLDID
 ECRTPENCRDMCMYONGYICITPRINRYGPNYSNPYSTGYPAAPLSLAP
 NYPTISRLICRFGIOMDESNOCDVDECTSHOCNPICISTEPGYSACSDGTV
 LLEGQCLDIDECRYGQQLCANVPGSYSCNPGFTLANDEGSCQOVNCACTENPV
 QCVNATGSPFICRDPGELEDEYVHSCSDMDECSFSEFLCOHECVNPGYFCSCP
 YLLDNRSCODINECEHNRHTCNLOOTCEHGFPCIDPICEEYELIISDRMC
 PLENDCRDPFTILYRDMVYSGRVPADIDFOMATTPRPAVYIYQISNGRRE
 YMRGTSPISATIVMRPIKRGPREIQLDLEKITVYINFGSSVIRLRLVTSQIPF"

CDS

BASE COUNT 588 a 650 c 547 g 582 t
 ORIGIN

Query Match 100.0%; Score 1344; DB 9; Length 2367;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGCCAGAAATTAAGAAGATCTACTGTTACCAATTCGTGCTCTTCCAAAGCCT 60
DB 184 ATGCCAGAAATTAAGAAGATCTACTGTTACCAATTCGTGCTCTTCCAAAGCCT 243
OY 61 GGAATGCACAGCAGCAGTGCAGCAAGTGGCTTGACCTGAGTGCAGTACAGTGT 120
DB 244 GGAATGCACAGCAGCAGTGCAGCAAGTGGCTTGACCTGAGTGCAGTACAGTGT 303
OY 121 TTAGATATTGATGAATGCGCAACCAATCCCGAGGCTCGCGAGGAGATGATGTGT 180
DB 304 TTAGATATTGATGAATGCGCAACCAATCCCGAGGCTCGCGAGGAGATGATGTGT 363
OY 181 AACCAAAATGGCGGGTATTTATGCAATTCGCCGACAAACCCGTGTATCGAGGGCCCTAC 240
DB 364 AACCAAAATGGCGGGTATTTATGCAATTCGCCGACAAACCCGTGTATCGAGGGCCCTAC 423
OY 241 TCGAACCCCTACTGACCCCTACTGACGTCAGGTCCGTACCCGACAGCTGCCCAACCACTCA 300
DB 424 TCGAACCCCTACTGACCCCTACTGACGTCAGGTCCGTACCCGACAGCTGCCCAACCACTCA 483
OY 301 GCTCCAAACTATCCACGATCTCCAGGCTCTTATATGCGCGTTTGATACCAAGATGAT 360
DB 484 GCTCCAAACTATCCACGATCTCCAGGCTCTTATATGCGCGTTTGATACCAAGATGAT 543
OY 361 GAAAGCAACCAATGTGTGATGTGACAGAGTGCACAAGATTCACCAAGTGCACACCC 420
DB 544 GAAAGCAACCAATGTGTGATGTGACAGAGTGCACAAGATTCACCAAGTGCACACCC 603
OY 421 ACCGAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCTGACACCAAGGATATG 480
DB 604 ACCGAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCTGACACCAAGGATATG 663

```

```

OY 481 CTTCTGGAAGGCCAGTGTCTAGACATGATGATGATGCTGATGTTACTGTCACAGCTC 540
DB 664 CTTCTGGAAGGCCAGTGTCTAGACATGATGATGATGCTGATGTTACTGTCACAGCTC 723
OY 541 TGTGCAAAATGTTCTGTGATCTTATTTCTGTATCATGACAAACCTGTGTTTAACTCAATGAG 600
DB 724 TGTGCAAAATGTTCTGTGATCTTATTTCTGTATCATGACAAACCTGTGTTTAACTCAATGAG 783
OY 601 GATGGAAGGCTTGTCCAAAGATGGAAGATGAGTGTGCGACCAAGAACCCCTGCTCAAAAC 660
DB 784 GATGGAAGGCTTGTCCAAAGATGGAAGATGAGTGTGCGACCAAGAACCCCTGCTCAAAAC 843
OY 661 TCGGTCAACACCTACAGGCTCTTTCATCTGCGCGCTGTGACCCAGAGATATGAATGAGAA 720
DB 844 TCGGTCAACACCTACAGGCTCTTTCATCTGCGCGCTGTGACCCAGAGATATGAATGAGAA 903
OY 721 GATGCGCTTCATTTGCAATGATATGACAGAGTGTGCACTTCTGTGAGTTCTCTGCCAATAT 780
DB 904 GATGCGCTTCATTTGCAATGATATGACAGAGTGTGCACTTCTGTGAGTTCTCTGCCAATAT 963
OY 781 GAGTGTGTGAACACAGCCCGGACATCTTCTGCTGCTGCGCTCCAGGCTACATCTGCTG 840
DB 964 GAGTGTGTGAACACAGCCCGGACATCTTCTGCTGCTGCGCTCCAGGCTACATCTGCTG 1023
OY 841 GATGACAAACCAAGCTGCGCAAGACATCAAGAAATGTGAGCAGAGAAACACAGCTGCAAC 900
DB 1024 GATGACAAACCAAGCTGCGCAAGACATCAAGAAATGTGAGCAGAGAAACACAGCTGCAAC 1083
OY 901 CTGCGACGAGAGCTGTACAAATTTACAAAGGGGCTTCAAAATGATTCGACCCCATCCGCTGT 960
DB 1084 CTGCGACGAGAGCTGTACAAATTTACAAAGGGGCTTCAAAATGATTCGACCCCATCCGCTGT 1143
OY 961 GAGGAGCCTTATCTGAGATCAGTGTATTAACCGCTGATGTGCTGCTGAGAACCTGAC 1020
DB 1144 GAGGAGCCTTATCTGAGATCAGTGTATTAACCGCTGATGTGCTGCTGAGAACCTGAC 1203
OY 1021 TCGAGAGACAGCCCTTTACCAATCTTTGTACCGGGACATGAGAGTGTGACAGAGCTGC 1080
DB 1204 TCGAGAGACAGCCCTTTACCAATCTTTGTACCGGGACATGAGAGTGTGACAGAGCTGC 1263
OY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCAGCCGCTACCGCTGGGGCTTATTCATT 1140
DB 1264 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCAGCCGCTACCGCTGGGGCTTATTCATT 1323
OY 1141 TTCCAGATCAAAATCTGGAATGAGGGCAGAGAAATTTACATGCGGCAAAAGGGCCCATC 1200
DB 1324 TTCCAGATCAAAATCTGGAATGAGGGCAGAGAAATTTACATGCGGCAAAAGGGCCCATC 1383
OY 1201 AGTGCACACCGGTGATGACACAGCCCATCAAGAGGGCCCGGGAAATCCAGCTGACCTTG 1260
DB 1384 AGTGCACACCGGTGATGATGACAGCCCATCAAGAGGGCCCGGGAAATCCAGCTGACCTTG 1443
OY 1361 GAAATGATCACTGTCAACACTGTCAATCAACTCAGAGGAGTCCGTGATCCGAGCTGGCG 1320
DB 1444 GAAATGATCACTGTCAACACTGTCAATCAACTCAGAGGAGTCCGTGATCCGAGCTGGCG 1503
OY 1321 ATATATGTGTGCAATGCCATTC 1344
DB 1504 ATATATGTGTGCAATGCCATTC 1527

```

RESULT 5
 LOCUS AR036548
 DEFINITION Sequence 2 from patent US 5872234.
 ACCESSION AR036548
 VERSION AR036548.1 GI:5953216
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2550)

source
1. .2328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
map:14q32.1
169.1515
/note="DANCE: contains 6 cBEGF domains and 1 Rgp motif;
secreted protein; expressed in developmental aorta, neural
crest cells, balloon injured vessels, atherosclerotic
lesions"
/codon_start=1
/product="developmental arteries and neural crest EGF-like
protein"
/protein_id="AAD1768.1"
/db_xref="GI:5305673"
/translation="MGKIRLTVTLALCLPSRGNQAOCTNFDLDROGCLDID
ERTIPACRDMCVNONGTYLIPITNFTVKGPISNPSTYSGPIPAAPLPSAP
NPPTISRLPREFGYOMDESNOQVDVDECAIDSHQNPOTICINTEGATSCIDGYW
LLEGQDLIDECRYGCOOLCANVPSSCTCNPFYLMEDGRSCODVNECATENPCV
QTCVNTYGSFICRDEYSELEEDGVHSCSDDECSFSEPLCOHECNODYFSCSPG
YLLDNRSCODINECEHRHNTCNLOOTCYNLOGKFCIDIRCEEPYLRISDNRCM
PAPNPGCRDOPFTILYRDMDVSGRSVPADIFOMATRYGAYIYIPIKSGNEREF
YMRQGPISATLMTPIKGPRIQDLEMTVTYVINFKSSVIRLRITYSQIF"

BASE COUNT 560 a 646 c 341 g 581 t

Query Match
Best Local Similarity 100.0%; Score 1344; DB 9; Length 2328;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 ATGCCAGAAATATAAGATATCTACTGTATACCATCTGCTCTGTCTTCCAGCCCT 60
169 ATGCCAGAAATATAAGATATCTACTGTATACCATCTGCTCTGTCTTCCAGCCCT 228
61 GGGAAATGCACAGGCAAGTGCAGAAATGCTTACCTGTGATCCGCACTGAGACAGT 120
229 GGGAAATGCACAGGCAAGTGCAGAAATGCTTACCTGTGATCCGCACTGAGACAGT 288
121 TTAGATATGATGAAATGCGCAACCATCCCGAGGCTGCGGAGAGACATATGCTGT 180
289 TTAGATATGATGAAATGCGCAACCATCCCGAGGCTGCGGAGAGACATATGCTGT 348
181 AAACCAAAATGGCGGGTATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTAC 240
349 AAACCAAAATGGCGGGTATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTAC 408
241 TGGAAACCCCTACTCGACACCCCTACTAGTCCGTCACCAAGAGTGCCTCCCACTCTCA 300
409 TGGAAACCCCTACTCGACACCCCTACTAGTCCGTCACCAAGAGTGCCTCCCACTCTCA 468
301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGGAT 360
469 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCGCTTGGATACAGATGGAT 528
361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACACAGATTTCCACACAGTCAACCC 420
529 GAAAGCAACCAATGTGTGATGTGAGAGAGTGCACACAGATTTCCACACAGTCAACCC 588
421 ACCCAAGATGTGATCAATATGAAAGGGGGGTACACCTGCTCTCGACGAGAGATTTGG 480
589 ACCCAAGATGTGATCAATATGAAAGGGGGGTACACCTGCTCTCGACGAGAGATTTGG 648
481 CTTCCTGGAAGGCGAGTGTAGACATGATGATGATGATGATGATGATGATGATGATG 540
649 CTTCCTGGAAGGCGAGTGTAGACATGATGATGATGATGATGATGATGATGATGATG 708
541 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
709 TGTGGAATGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 768
601 GATGGAAGGTCTTGGCAAGATGTGCAAGAGTGTGCGCAACCGGAGAACCCCTGGTGAAC 660
769 GATGGAAGGTCTTGGCAAGATGTGCAAGAGTGTGCGCAACCGGAGAACCCCTGGTGAAC 828

Query 661 TGGCTCAACACCTACAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
829 TGGCTCAACACCTACAGGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 888
721 GATGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
889 GATGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 948
781 GATGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
949 GATGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008
841 GATGACACACCAAGGCTCCCAAGACATCAACGAATGTGAGCAGAGAACACAGTCAAC 900
1009 GATGACACACCAAGGCTCCCAAGACATCAACGAATGTGAGCAGAGAACACAGTCAAC 1068
901 CTGACACACAGAGTGTCTCAATTTACAAAGGGGGCTTCAATGATGATGATGATGATG 960
1069 CTGACACACAGAGTGTCTCAATTTACAAAGGGGGCTTCAATGATGATGATGATGATG 1128
961 GAGAGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
1129 GAGAGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1188
1021 TGCAGAGACACAGGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
1189 TGCAGAGACACAGGCTTTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1248
1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACAGCCCTACCTCGGGGCTTATTAAT 1140
1249 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACAGCCCTACCTCGGGGCTTATTAAT 1308
1141 TTCCAGATCAATCTGGGAATGAGGCAAGAAATTTTCAATGCGGCAAGGGGCCCATC 1200
1309 TTCCAGATCAATCTGGGAATGAGGCAAGAAATTTTCAATGCGGCAAGGGGCCCATC 1368
1201 AGTCCACACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
1369 AGTCCACACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1428
1261 GAAATGATCACTGTCAACACTGTGATGATGATGATGATGATGATGATGATGATGATG 1320
1429 GAAATGATCACTGTCAACACTGTGATGATGATGATGATGATGATGATGATGATGATG 1488
1321 ATATATGTGTGCGAGTACCATTC 1344
1489 ATATATGTGTGCGAGTACCATTC 1512

RESULT 4
LOCUS BC022280
DEFINITION Homo sapiens, clone MGC:22412 IMAGE:4693953, mRNA, complete cds.
ACCESSION BC022280
VERSION BC022280.1 GI:18490144
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2367)
AUTHORS Strausberg, R.
TITLE Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDS

1. .1347
/gene="FIBL-5"
/codon_start=1
/product="Fibulin-5"
/protein_id="CAB38568.1"
/db_xref="GI:4490530"
/translation="MPEIKRILVTILALCLPSRNAOAOCTNGEDFDIDROSGGOCID
ERTIPEACRGMVCNONGVYICIPRTNPNYRPNSTPYSGLPAPAAPLSAP
NYPTISPLICRFGYQDNESNOCVDDECATDSDHOCNPOTICITNEGTCSDGW
LLEAGTDLIDECRYGCOQLCANVPSSYSCCTNPGFTLNEBRSODVNECATENPCV
QCVNTGYSEICRDPGELEEDGVHCSDECSFSEFLCHECVNPGFTYSCSPG
YLIDNRSODINEEHRNHTNLOQTQCYNLOGFKCIDIRCEEPYLRISDNCMC
PAMPGRDOPFTILYRDMDVSVSRPADIFOMATRYGAYIYDRIKSNMCR
VMRNGPISATLVMTRPIKGRIRIQLDLEMTVNTVINFRRSSVIRLRITYSQIP"

sig_peptide
1. .66
/gene="FIBL-5"
/gene="FIBL-5"
67. .1344
/product="Fibulin-5"

BASE COUNT 507 a 540 c 475 g 496 t 1 others

Query Match 100.0%; Score 1344; DB 9; Length 2019;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGGAATAAAGATACCTACCTTACCATTTCTGCTCTCTTCCAAAGCCCT 60
1 ATGCCAGGAATAAAGATACCTACCTTACCATTTCTGCTCTCTTCCAAAGCCCT 60
61 GGGAAATGCACAGGACAGTGCAGAAATGCTTGGATGATCGCCAGTGCAGAGTGT 120
61 GGGAAATGCACAGGACAGTGCAGAAATGCTTGGATGATCGCCAGTGCAGAGTGT 120
61 GGGAAATGCACAGGACAGTGCAGAAATGCTTGGATGATCGCCAGTGCAGAGTGT 120
121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
121 TTAGATATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
181 AACCAAAATGGGGGATTTATGATTCCTCCGGAGCAAAACCTGTGATTCGAGGGCCCTAC 240
181 AACCAAAATGGGGGATTTATGATTCCTCCGGAGCAAAACCTGTGATTCGAGGGCCCTAC 240
181 AACCAAAATGGGGGATTTATGATTCCTCCGGAGCAAAACCTGTGATTCGAGGGCCCTAC 240
241 TCGAACCCCTACTCGACCCCTACTAGTCCGTACCCAGCAGCTGCCACACACTCTCA 300
241 TCGAACCCCTACTCGACCCCTACTAGTCCGTACCCAGCAGCTGCCACACACTCTCA 300
301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCGCTTGGATACCAAGATGAT 360
301 GCTCCAAATATCCCAAGATCTCCAGGCTCTTATATGCCGCTTGGATACCAAGATGAT 360
361 GAAAGCAACCAATGTGTGATGAGGAGTGTGCAAAAGATTCGCCACAGTGTGCAACCC 420
361 GAAAGCAACCAATGTGTGATGAGGAGTGTGCAAAAGATTCGCCACAGTGTGCAACCC 420
421 ACCGATCTGCATCAATCTAGTGAAGGCGGTACACCTGTCTCTGACCGAGATATGG 480
421 ACCGATCTGCATCAATCTAGTGAAGGCGGTACACCTGTCTCTGACCGAGATATGG 480
421 ACCGATCTGCATCAATCTAGTGAAGGCGGTACACCTGTCTCTGACCGAGATATGG 480
481 CTTCTGGAAGGCGATGCTTAGACATGTGAATGTGCTGATGATGATGATGATGATGAT 540
481 CTTCTGGAAGGCGATGCTTAGACATGTGAATGTGCTGATGATGATGATGATGATGAT 540
541 TGTGGAATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
541 TGTGGAATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
601 GATGGAAGTCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601 GATGGAAGTCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
661 TGCCTCAACACCTACCGCTCTTCACTGCTCCGCTGTGACCCAGATATGAATGAGGAA 720
661 TGCCTCAACACCTACCGCTCTTCACTGCTCCGCTGTGACCCAGATATGAATGAGGAA 720

721 GATGCGCTTCATTCAGTATGATGAGGAGTGCAGCTTCTCTGAGTCTCTCTGCCAAT 780
721 GATGCGCTTCATTCAGTATGATGAGGAGTGCAGCTTCTCTGAGTCTCTCTGCCAAT 780
781 GATGCTGTGAACAGCCCGGACACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
781 GATGCTGTGAACAGCCCGGACACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
841 GATGCTGTGAACAGCCCGGACACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
841 GATGCTGTGAACAGCCCGGACACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
841 GATGCTGTGAACAGCCCGGACACTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
901 CTGACAGACAGTGTCTCAATTTAACAAGGGGCTTCAAAATGATGATGATGATGATGAT 960
901 CTGACAGACAGTGTCTCAATTTAACAAGGGGCTTCAAAATGATGATGATGATGATGAT 960
961 GAGGAGCCCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 GAGGAGCCCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 TCAGAGACAGGAGCCCTTACATCTTGTACCGGAGATGAGAGTGTCTGAGAGCTTC 1080
1021 TCAGAGACAGGAGCCCTTACATCTTGTACCGGAGATGAGAGTGTCTGAGAGCTTC 1080
1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCTGCGGCTATTTACAT 1140
1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCCGCTACCTGCGGCTATTTACAT 1140
1141 TTCCAGATCAAAATCTGGGAATGAGGGGAGGAATTTAACAATGGGGCAAGGGGCCCTAC 1200
1141 TTCCAGATCAAAATCTGGGAATGAGGGGAGGAATTTAACAATGGGGCAAGGGGCCCTAC 1200
1201 ACTGCCACCTGTGTGATGACACGCCCCATCAAAAGGGCCCGGGAAATCCAGCTGACTTC 1260
1201 ACTGCCACCTGTGTGATGACACGCCCCATCAAAAGGGCCCGGGAAATCCAGCTGACTTC 1260
1261 GAAATGATATCTGTCAACACTGTCTATCACTTACAGAGGAGCTCTGATGATGATGATGAT 1320
1261 GAAATGATATCTGTCAACACTGTCTATCACTTACAGAGGAGCTCTGATGATGATGATGAT 1320
1321 ATATATGTGTGAGTACCATTC 1344
1321 ATATATGTGTGAGTACCATTC 1344

RESULT 3
AF112152 2328 bp mRNA linear PRI 02-AUG-1999
LOCUS AF112152
DEFINITION Homo sapiens developmental arteries and neural crest EGF-like
ACCESSION AF112152
VERSION AF112152.1 GI:5305672
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2328)
Nakamura, Y., Kohno, K., Tashiro, K., Ito, Z., Ando, N., Ito, S., Schaub, R.,
Matsunaga, A., Saegusa, S., Giken, K., and Honjo, T.
DAB2, a novel secreted EGF protein expressed in developing,
embryonic, and balloon-injured arteries
JOURNAL OF BIOLOGICAL CHEMISTRY 274 (32): 22476-22483 (1999)
MEDLINE 958567479
PUBMED 10428823
REFERENCE 2 (bases 1 to 2328)
Nakamura, Y., Tashiro, K., and Honjo, T.
Direct Submission
Submitted (09-DEC-1998) Medicine, University of California San
Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
FEATURES
Location/Qualifiers

FEATURES

URES source	Location/Qualifiers
	1. .1720
	/organism="unidentified"
	/db_xref="taxon:32644"
CDS	211. .1557

	a	c	g	t	others
sig_peptide	211	.284			
mat_peptide	287	.1554			
BASE COUNT	396 a	515 c	434 g	372 t	3 others
ORIGIN					

Query Match	100.0%	Score 1344:	DB 6:	Length 1720:
Best local similarity	100.0%	Pred. No. 0:		
Matches 1344:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:

OY	1	ATGCCAGGAATTA	AAAAGGATACACGTTTACCATCTTGCGTCTCTGCTCCAAAGCCCT	60
Db	211	ATGCCAGGAATTA	AAAAGGATACACGTTTACCATCTTGCGTCTCTGCTCCAAAGCCCT	270
OY	61	GGAATGACACAGG	CACAGTGCAGGAATGGCTTTGACCTGATGCCAGTGCAGCAGTGT	120
Db	271	GGAATGACACAGG	CACAGTGCAGGAATGGCTTTGACCTGATGCCAGTGCAGCAGTGT	330
OY	121	TTTGATATTGAT	ATGAATGCCGAACCATGCCGAGGCGCCGAGAGACATGATGTGT	180
Db	331	TTTGATATTGAT	ATGAATGCCGAACCATGCCGAGGCGCCGAGAGACATGATGTGT	390
OY	181	AACCAAAATGG	CGGGATATTTATGCATTTCCCGGCAAAACCTGTGTATCGAGGGCCCTAC	240
Db	391	AACCAAAATGG	CGGGATATTTATGCATTTCCCGGCAAAACCTGTGTATCGAGGGCCCTAC	450
OY	241	TGCACCCCTAC	TCGACCCCTATCTACAGGTCGACACAGAGTGGCCCAACACTCTCA	300
Db	451	TGCACCCCTAC	TCGACCCCTATCTACAGGTCGACACAGAGTGGCCCAACACTCTCA	510
OY	301	GCTCAAACTAT	CCACGATCTCCAGGCTCTTATATGCCGCTTGTGATACCAATGGAT	360
Db	511	GCTCAAACTAT	CCACGATCTCCAGGCTCTTATATGCCGCTTGTGATACCAATGGAT	570
OY	361	GAAAGCAACCA	ATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACCACTGCACATCCC	420
Db	571	GAAAGCAACCA	ATGTGTGATGTGAGAGAGTGTGCACAGATTTCCACCACTGCACATCCC	630
OY	421	ACCAGATCTGC	ATCACTACTGAAAGGGGGGTACACTGCTCTGCACGAGGATATTTGG	480
Db	631	ACCAGATCTGC	ATCACTACTGAAAGGGGGGTACACTGCTCTGCACGAGGATATTTGG	690
OY	481	CTTCTGGAAG	CGCAGTCTTAGACATTGATGATGTGCTATGTACTGCCAGCAGCTC	540
Db	691	CTTCTGGAAG	CGCAGTCTTAGACATTGATGATGTGCTATGTACTGCCAGCAGCTC	750
OY	541	TGTGCGAATTC	CTCGATCTATCTTTGACATGCAACCTGGTTTAACTCCGTAAGAG	600
Db	751	TGTGCGAATTC	CTCGATCTATCTTTGACATGCAACCTGGTTTAACTCCGTAAGAG	810
OY	601	GATGGAAGGCT	TTGCCAAGATGTGAACGAGTGTCCACCGAACCCTCGCTGCACAAAC	660
Db	811	GATGGAAGGCT	TTGCCAAGATGTGAACGAGTGTCCACCGAACCCTCGCTGCACAAAC	870
OY	661	TGCGTCAACAC	CTAAGGCTCTTTTCACTGCGCGCTGTACCCAGATATGAACCTTGAGAA	720
Db	871	TGCGTCAACAC	CTAAGGCTCTTTTCACTGCGCGCTGTACCCAGATATGAACCTTGAGAA	930

OY		721	GATGGGCTCATTTGAGTGAATTATGGACGACTGCAGCTTCCTGAGTTCCTGTCCGCAACAT	780
Db		931	GATGGGCTTCATTGCACTGATTAATGAGCAGCTGCAGCTTCCTGAGTTCCTGTCCCAACAT	990
OY		781	GACTGTGTGAACAAGCGCCGCCACATACTTGCTGCTCCGCCCTCCAGGCTCTACATCCTGCTG	840
Db		991	GACTGTGTGAACAAGCGCCGCCACATACTTGCTGCTCCGCCCTCCAGGCTCTACATCCTGCTG	1050
OY		841	GATGACAAACCGAAGCTGCCCAAGACATCACGAATGTGACACAGAAACCACAGCTGCAC	900
Db		1051	GATGACAAACCGAAGCTGCCCAAGACATCACGAATGTGACACAGAAACCACAGCTGCAC	1110
OY		901	CTGCAGACAGACGTCATCAATTTTACAAGGGGGGCTTCAAATGATCGAACCCCATCGCTGT	960
Db		1111	CTGCAGACAGACGTCATCAATTTTACAAGGGGGGCTTCAAATGATCGAACCCCATCGCTGT	1170
OY		961	GAGAGAGCTTATCTGAGGATCAGTGATTAACCGCTGATGTGTCTGCTGTGAGAACCTTGCC	1020
Db		1171	GAGAGAGCTTATCTGAGGATCAGTGATTAACCGCTGATGTGTCTGCTGTGAGAACCTTGCC	1230
OY		1021	TGCAGAGACCAAGCCCTTTACCATTCTTGTACCGGGACATGACACGTGGTGTACGAGCGCTCC	1080
Db		1231	TGCAGAGACCAAGCCCTTTACCATTCTTGTACCGGGACATGACACGTGGTGTACGAGCGCTCC	1290
OY		1081	GTTCGCCGCTGACATCTTCCAAATGACAGCAGCCGCTACACCTCGGGGGCTATTACATT	1140
Db		1291	GTTCGCCGCTGACATCTTCCAAATGACAGCAGCCGCTACACCTCGGGGGCTATTACATT	1350
OY		1141	TTCGACATCAAAATCTGGGAATGAGGGGACAGAAATTTTACATCGGGCAAAACGGGGCCCATC	1200
Db		1351	TTCGACATCAAAATCTGGGAATGAGGGGACAGAAATTTTACATCGGGCAAAACGGGGCCCATC	1410
OY		1201	AGTGCACCCCTGGTGATGACACAGCCCCCATCAAGAAGGGCCCGGGAAATCCAGCTGACCTG	1260
Db		1411	AGTGCACCCCTGGTGATGACACAGCCCCCATCAAGAAGGGCCCGGGAAATCCAGCTGACCTG	1470
OY		1261	GAATATGATCACTGTCACACACTGTCATCACTCAACTTCAGAGGACAGCTCCGATCCGACTGCGG	1320
Db		1471	GAATATGATCACTGTCACACACTGTCATCACTCAACTTCAGAGGACAGCTCCGATCCGACTGCGG	1530
OY		1321	ATATATGTGTGCGCAGTAGTACCATTTC	1344
Db		1531	ATATATGTGTGCGCAGTAGTACCATTTC	1554
RESULT 2				
HSAL133490				
LOCUS	HSAL133490	2019 bp	mRNA	linear PRI 11-MAR-1999
DEFINITION	Homo sapiens fibulin-5.			
ACCESSION	AJ133490			
VERSION	AJ133490.1 GI:4490529			
KEYWORDS	FTBL-5 gene; fibulin-5.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
JOURNAL	1 (bases 1 to 2019)			
	Kostka,G.			
	Direct Submission			
	Submitted (05-MAR-1999) Kostka G., Dept. of Protein Chemistry,			
	Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 82152			
	Marlinsried, GERMANY			
REFERENCE	2 (bases 1 to 2019)			
AUTHORS	Kostka,G.			
JOURNAL	Unpublished			
FEATURES	Location/Qualifiers			
source	1..2019			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/tissue_type="melanoma"			
gene	1..1347			
	/gene="FIBL-5"			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 3504.57 Seconds

(without alignments)
11160.894 Million cell updates/sec

Files: US-09-674-379a-11

Sequence: 1 atgcgcagcaataaaagat.....atgtctgcagtaaccattc 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank
1: gb-bank
2: gb-bank
3: gb-in
4: gb-om
5: gb-ov
6: gb-pat
7: gb-ph
8: gb-pl
9: gb-pr
10: gb-ro
11: gb-sts
12: gb-sy
13: gb-un
14: gb-vl
15: em-ba
16: em-fun
17: em-hum
18: em-in
19: em-mu
20: em-om
21: em-or
22: em-ov
23: em-ph
24: em-pl
25: em-ro
26: em-sts
27: em-un
28: em-vl
29: em-vl
30: em-hg-hum
31: em-hg-inv
32: em-hg-other
33: em-hg-mus
34: em-hg-pln
35: em-hg-rod
36: em-hg-mam
37: em-hg-vrt
38: em-sy
39: em-hggo-hum
40: em-hggo-mus
41: em-hggo-other

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1720	6 AB4086	AB4086 Sequence 1
2	1344	100.0	2019	9 HSA133490	AJ133490 Homo sapi
3	1344	100.0	2338	9 AF112152	AF112152 Homo sapi
4	1344	100.0	2367	9 BC022280	BC022280 Homo sapi
5	1344	100.0	2550	6 AR036548	AR036548 Homo sapi
6	1344	100.0	2550	6 AR173204	AR173204 Sequence
7	1342.4	99.9	2609	6 AX376340	AX376340 Sequence
8	1342.4	99.9	2609	6 AX403659	AX403659 Sequence
9	1338	98.8	2019	9 AF093118	AF093118 Homo sapi
10	1113.6	82.9	2230	10 BC006636	BC006636 Mus muscu
11	1113.6	82.9	2478	10 AF112151	AF112151 Mus muscu
12	1104	82.1	2499	10 AF112153	AF112153 Rattus no
13	1102.4	82.0	2304	10 AF137350	AF137350 Rattus no
14	372.8	27.7	1440	9 HSA132819	AJ132819 Homo sapi
15	372.8	27.7	1536	9 AB030655	AB030655 Homo sapi
16	371.2	27.6	1480	6 AX023967	AX023967 Sequence
17	371.2	27.6	1561	6 BC010456	BC010456 Homo sapi
18	371.2	27.6	1707	9 AK000980	AK000980 Homo sapi
19	371.2	27.6	1875	6 AX201326	AX201326 Sequence
20	371.2	27.6	1875	6 AX464184	AX464184 Sequence
21	371.2	27.6	2018	6 AR095382	AR095382 Sequence
22	369.6	27.5	1757	9 AF109121	AF109121 Homo sapi
23	364.8	27.1	1825	9 AF124486	AF124486 Homo sapi
24	363.8	27.1	1705	9 AF093119	AF093119 Homo sapi
25	350.4	26.1	1778	10 AF046870	AF046870 Cricetus
26	347.2	25.8	1512	10 AF104223	AF104223 Mus muscu
27	347.2	25.8	1513	6 AX023961	AX023961 Sequence
28	347.2	25.8	1794	10 AF109122	AF109122 Mus muscu
29	345.6	25.7	1724	10 BC012269	BC012269 Mus muscu
30	337.6	25.1	1358	6 AX023976	AX023976 Sequence
31	312	23.2	1021	6 AX023954	AX023954 Sequence
32	293.8	22.0	2613	9 AK098186	AK098186 Homo sapi
33	268	19.9	1122	6 AX023965	AX023965 Sequence
34	260.6	19.4	171901	9 AC007057	AC007057 Homo sapi
35	260.6	19.4	172674	9 CENS0008	AL049872 Human chr
36	260.6	19.4	175210	9 CENS07EG5	AL590328 Human chr
37	250.6	18.6	1677	10 BC031184	BC031184 Mus muscu
38	249	18.5	1911	10 DB9730	DB9730 Rattus rat
39	241	17.9	2025	9 BC014410	BC014410 Homo sapi
40	241	17.9	2512	6 AX277602	AX277602 Sequence
41	241	17.9	2512	6 AX331916	AX331916 Sequence
42	241	17.9	2512	6 AX332783	AX332783 Sequence
43	241	17.9	2512	6 AX333462	AX333462 Sequence
44	241	17.9	2512	6 AX334103	AX334103 Sequence
45	241	17.9	2512	6 AX335899	AX335899 Sequence

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
AB4086	AB4086	Sequence 1 from Patent WO9846746.	AB4086	AB4086.1	GI:6733224	unclassified.	unclassified.	1 (bases 1 to 1720)	Li, H. and Olsen, H. S.	EXTRACELLULAR/EPIDERMAL GROWTH FACTOR LIKE PROTEIN	Reference WO 9846746-2 A 1 22-06-1998
											HUMAN GENOME SCIENCES INC (US) R. L. HAO DONG (US)

Db 93 AGATCCCTACATCTTAACACAGAAACCGATGCTTTGCCAGCTCTCAATGCCATGTG 152
QY 1023 CAGAGACCCCTTTTACATCTTGTACCGGGAGTACGAGCGTGTGAGAGCCGCT 1082
Db 153 CCGAGAACTCCCGACGATAGTCTTCAATATACATGACATCCATGATAGGCTCT 212
QY 1083 TCCCGGTGACATCTTCCAAATGACAGCAGCCGCTACCTGGGCTTATACATTT 1142
Db 213 GCCATGACATCTTCCAGATACAGGCCACAACTATTTATGCAACACATCATATCTT 272
QY 1143 CCAGATCAATCTGCGAATGAGGCGACAGAAATTTTACATGCGGCAAGGCGCCCTG 1202
Db 273 TCGGATTAACTGGAATGAAATGAG-AGTCTACCTACGCAACAANCCCTGTAG 331
QY 1203 TGCACCCCTGCTGATGACAGCCCATCAAGGCGCCGGAATCCAGCTGACTTGA 1262
Db 332 TGCATGCTGTCTGCTGCTGAAGNCATTTATCAGGACCAAGAACATATCTGACTGA 391
QY 1263 AATGATCAC 1271
Db 392 GATGCTGAC 400

RESULT 10
US-09-404-879a-340/c
; Sequence 340, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 340
; LENGTH: 220
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(220)
; OTHER INFORMATION: n - A,T,C or G
US-09-404-879a-340

Query Match 6.0%; Score 80.2; DB 4; Length 220;
Best Local Similarity 60.5%; Pred. No. 5.5e-16;
Matches 130; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 373 TGTGTGATGTGACGAGTGTGCAACAGATTTCCACAGTGCACCCACCAATCTGC 432
Db 219 TGTGTGATGTGACGAGTGTGCAACAGTGTGCAACAGTGTGCAACAGTGTGCA 160
QY 433 ATCAATAGTGAAGGCGGTACACCTGCTCTGCAACGCGGATATTTGCTTGAAGGC 492
Db 159 CATAACTGCTGCTGCTCTCTCACTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 100
QY 493 CAGTGTGATGACATGATGATGATGCTGCTGATGATGATGATGATGATGATGATG 552
Db 99 GAGTGTGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 40
QY 553 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 587
Db 39 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5

RESULT 11
US-08-282-141-1
; Sequence 1, Application US/08282141
; Patent No. 5538861

GENERAL INFORMATION:
; APPLICANT: Schneider, Claudio
; APPLICANT: Varnum, Brian
; APPLICANT: Avanzi, Giancarlo
; APPLICANT: Brancolini, Claudio
; APPLICANT: Manioldi, Guido
; TITLE OF INVENTION: Stimulating Factor for the AXL Receptor
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehaven Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: United States
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,141
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2461 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-08-282-141-1

Query Match 4.3%; Score 58.2; DB 1; Length 2461;
Best Local Similarity 54.5%; Pred. No. 2.7e-08;
Matches 164; Conservative 0; Mismatches 128; Indels 9; Gaps 2;

QY 521 ATGTTACTGCGACAGCTGTGCGAATGCTGATGATGATGATGATGATGATGATGATG 580
Db 625 ACGGGGCTGCTCTCAAGATCTGCGCAACAGCCGCGTACCTTCTTCTGCGACA 684
QY 581 CTGTTTACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640
Db 685 GCGCTTGCAGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 744
QY 641 AGAACCTCTGCTGCAACCTGCTGCAACCTA---CGCTCTTCACTGCTGCTG 697
Db 745 CCGAGGCTGCGGGGAGGCGGCTGCAAGAACCTGCGGCTCTCTCTGCTG 804
QY 698 ACCGAGATATGACTGAGAGATGCGCTTCACTGATGATGATGATGATGATGATG 757
Db 805 ACGAGGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 859
QY 758 TCTGTGATGCTTCTGCGAATGATGATGATGATGATGATGATGATGATGATGATG 817
Db 860 -TCTGAGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 918
QY 818 G 818
Db 919 G 919

RESULT 12
US-08-479-722B-3
; Sequence 3, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF (BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:

; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James Greg
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods For Their Use
 ; FILE REFERENCE: 11000.101c1
 ; CURRENT APPLICATION NUMBER: US/09/188.930A
 ; CURRENT FILING DATE: 1998-11-09
 ; NUMBER OF SEQ ID NOS: 348
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO: 259
 ; LENGTH: 1018
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-188-930-259

Query Match 23.68; Score 316.6; DB 3; Length 1018;
 Best Local Similarity 62.38; Pred. No. 2.9e-91;
 Matches 515; Conservative 0; Mismatches 309; Indels 3; Gaps 1;

QY 518 GCTATGTTACTGTCAGACGCTCTGTGGAGTGTCCGAGATCCATTTCTGTACATGCA 577
 DB 1 GCTACCGCTACTGTCAGACGCTCTGTGGAGTGTCCGAGATCCATTTCTGTACATGCA 60
 QY 578 ACCCTGTTTACCTTCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 637
 DB 61 AGCCGGGCTTCCAGCTGGGGCTTAAACACCGCTCTCTGTGTGATGATGATGATGATGAT 120
 QY 638 CCGAGAACCCCTGCTGCTCAACCTGCTGCTCAACCTGCTGCTCAACCTGCTGCTGCTGCT 697
 DB 121 TGGGGGGCCCTGAG 180
 QY 698 ACCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 757
 DB 181 ACCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 758 TCTCTAGTCT 817
 DB 241 ACTCCAGATCT 300
 QY 818 GCCCTCAGAGCTACATCT 877
 DB 301 GCCCAGAGGTTTACAGCTGCTG---GCCACACGCTCTGCTCCACACATTTGATGATGATG 357
 QY 878 AGCAGAGAGACACACAGCTGCAACCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 937
 DB 358 AGCTGCTGAG 417
 QY 938 AATGATGACAGCCGCT 997
 DB 418 GCTGCTGAG 477
 QY 998 TGTGTCTGCTGAGAACCTGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1057
 DB 478 TCTGCGCGGCTCTCAACCTCTATGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
 QY 1058 TGGAGCTGTGTGAG 1117
 DB 538 TGACCATACCTCTGAG 597
 QY 1118 GCTACCTGAG 1177
 DB 598 TCTACCCCGGCTGCTCAACCT 657
 QY 1178 ACATGAG 1237
 DB 658 ACATTTAG 717
 QY 1238 CCGGAG 1297
 DB 718 CCGGAG 777
 QY 1298 GCAAGCTCTGATGAG 1344

DB 778 CCAGCTCTGATGAGGCTCACCGCTTGTAGGGGCTTACACCTTC 824

RESULT 7
 US-09-248-757-1
 ; Sequence 1, Application US/09248757
 ; Patent No. 6417342
 ; GENERAL INFORMATION:

; APPLICANT: STONE, EDWIN M.
 ; APPLICANT: SHEFFIELD, VAL C.
 ; TITLE OF INVENTION: MACULAR DEGENERATION DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: UIA-018.02
 ; CURRENT APPLICATION NUMBER: US/09/248.757
 ; CURRENT FILING DATE: 1999-02-12
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 1
 ; LENGTH: 2512
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-248-757-1

Query Match 17.98; Score 241; DB 4; Length 2512;
 Best Local Similarity 54.68; Pred. No. 8.1e-67;
 Matches 547; Conservative 0; Mismatches 445; Indels 9; Gaps 3;

QY 346 GGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 405
 DB 403 GGTACAGAGCAAGTGAACACACAGCTGTCCAGACATGACAGAGTGCATCGACGAGAG 462
 QY 406 CACCAAGTCAACCCACACAGATCTGATCAATCAATCAATCAATCAATCAATCAATCAAT 465
 DB 463 CACAAGTCAAG 522
 QY 466 ACCAGAGATTTGGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
 DB 523 CTTCTGATATGAG 582
 QY 523 GGTATGAG 582
 DB 583 GGTATGAG 642
 QY 583 GGTATGAG 642
 DB 643 GGTATGAG 702
 QY 643 AATCAATGCTGCTGAG 762
 DB 703 GATATGAG 762
 QY 703 GATATGAG 762
 DB 763 GATATGAG 822
 QY 763 GATATGAG 822
 DB 823 GATATGAG 882
 QY 823 GATATGAG 882
 DB 883 GATATGAG 942
 QY 883 GATATGAG 942
 DB 940 GATATGAG 996
 QY 943 GATATGAG 1002
 DB 997 GATATGAG 1056
 QY 1003 GATATGAG 1062
 DB 1057 GATATGAG 1116

QY 1117 CGCTACCCCTGGGCGCTATTACATTTTCAGATCAATCTGGGAATGAGGCGACAGAAATTT 1176
Db 1226 GTCTACCCCGGCTGTACATATGCTTTTCAGATCCGTGTGGAATCTGCAAGGGGAGCTTT 1285
QY 1177 TACATGCGGCAAAACGGGCGCCATCAGTGGCCACCTGGTATGACAGCCCATCAAAAGG 1236
Db 1286 TACATTAGCAAAATCAACAACTCAAGCCCATGTGCTCTCGCCCGCGGTGACAGGGC 1345
QY 1227 CCGCGGGAATCCACCTGACCTTGGAATATGATCATCTCAACACTGTCAACTTCAGA 1296
Db 1346 CCGCGGAGTACGTGCTGACCTGAGATGTCACCATGAATTCCTCATGTAGACTACCGG 1405
QY 1297 GCGAGCTCCGTGATCCGACCTGGAGATATGTGTGCGAGTACCATTC 1344
Db 1406 GCGAGCTGTGTACTAGGCTACCGCTCTTTGTAGGGGCTACACCTTC 1453

RESULT 5

US-08-980-514-2
; Sequence 2, Application US/08980514
; Patent No. 6004753
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN SI-5-ECMP-LIKE PROT
; TITLE OF INVENTION: EIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,514
; FILING DATE: Filed Herewith
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0436 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2018 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRESTNOT13
; CLONE: 2786449
; US-08-980-514-2

Query Match 27.6%; Score 371.2; DB 3; Length 2018;
Best Local Similarity 61.2%; Pred. No. 1.5e-108;
Matches 617; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

QY 337 TCCCGCTTTGATACCATGATGTAAGCAACCAATGTGTGATGTGAGAGTGCA 396
Db 530 TCCCGACCGAGCTATGAGCCCGACGATCAAGACAGCTGTGTGATGTGAGAGTGCC 589

QY 397 ACAGATTCACACAGGATGCAACCCACCCAGATGTGCATCAATACTGAAAGCGGGTACACC 456
Db 590 CAGGCGCCGACAGACTGTGCGCCCGACGAGAGCTGCATTAATGCTGTGCTTCATCAG 649
QY 457 TGCCTCCGACCGAGGATATGCTTCTGGAAGCGCAGTGTGTAGACTTATGAAATGT 516
Db 650 TGCACCTGCTGTATGTTTACCGCAAGATCGGGCCGAGTGTGTGAGCATAGACAGTGC 709
QY 517 CGCTATGTTTACTCCGACAGCTGTGTGCAATGTTCCTGATTCCTATTCTGTACATGC 576
Db 710 CGCTACCGCTATGCTGCAGACACCGCTGCTGAACCTGCTGCTCTCTCCGTCGACGTGC 769
QY 577 AACCTGTTTACCTCAATGAGAGTGAAGTCTTCCAAAGATGTAACGAGTGTGC 636
Db 770 GAGCGGGGCTTCCAGCTGGGGCTTAACAAACGCTCTCTGTGTGATGTGAGAGTGTGC 829
QY 637 ACCGAGAACCCCTGTCGCAACCTGGGTCAACCTACGCTCTTATCATGCGCGTGT 696
Db 830 ATGGGGGCCCATGCGACGACGCGCTTCAACTCTATGAGACCTTCTGTGTGCTGC 889
QY 697 GACCGAGATATGAATTTGAGGAAGATGCGCTTCACTTATGATGAGAGTGCAGC 756
Db 890 CACGAGGGTATGATGATGCAATCGGGATGCTTCTCTGCACTGATGATGATGATGATG 949
QY 757 TTCTGTAGTTCCTCTGCAATGAGTGTGTGAACAGCCCGACATTAATCTGTGCTGC 816
Db 950 TACTCAGCTTACCTCTGTGATGACGCTGCTGCTCAACGAGCGCGTCTCTGCTGCAC 1009
QY 817 TGCCCTCCAGGCTACATCATCTGTGATGACACGAGGCTGCCAAGACATCAAGATGT 876
Db 1010 TGCCCGACAGGTTATCCAGCTGCTG--GCCACAGCGCTCTGCGAAGCATTTGATGAGT 1066
QY 877 GAGCAGAGAACACAGCTGCAACCTGACAGACGAGCTGTCAATTTTACAGGGGGCTTC 936
Db 1067 GAGCTGTGTGCGACAGATGCTCCGAGGCCCAACCTGTGTCACTTCCATGAGGGGCTAC 1126
QY 937 AATGATGACACCCATCCGCTGTGAGAGGCTTATGTGAGATCACTGATTAACCGCTGT 996
Db 1127 CGCTGTGTGACACCAACCGCTGTGAGACCCCTTCACTCAAGGTCTGTGAGAACCGCTGT 1186
QY 997 ATGTGTCTGCTGAGAACCTGTGCTGACAGACGAGCCCTTTTACATCTTACCGGGAG 1056
Db 1187 CTCTGCGCGGCCCTCCACACCTCTATGTCGAGAGCAGCTTCAATGTTGACCGCTAC 1246
QY 1057 ATGAGAGTGTGTGAGAGCGCTCCGCTGACATCTTCCAAATGACAGCCAGCAGCC 1116
Db 1247 ATGACCATCACTCGGAGGCGAGCGCTGACGCTGTCCAGATCCAGGCGACCTGC 1306
QY 1117 CGCTACCCCTGGGCGCTATTACATTTTCAGATCAATCTGGGAATGAGGCGACAGAAATTT 1176
Db 1307 GTCTACCCCGGCTGTCAATATGCTTTTCAGATCCGTGTGGAATCTGCAAGGGGAGCTTT 1366
QY 1177 TACATGCGGCAAAACGGGCGCCATCAGTGGCCACCTGGTATGACAGCCCATCAAAAGG 1236
Db 1367 TACATTAGCAAAATCAACAACTCAAGCCCATGTGCTCTCGCCCGCGGTGACAGGGC 1426
QY 1237 CCGCGGGAATCCACCTGACCTTGGAATATGATCACTGTCAACACTGTCAACTTCAGA 1296
Db 1427 CCGCGGAGTACGTGCTGAGACCTGTGAGAGTGCATGCAATTCCTCATGAGACTACCGG 1486
QY 1297 GCGAGCTCCGTGATCCGAGCTGCGGAGATATGTGTGCGAGTACCATTC 1344
Db 1487 GCGAGCTGTGTACTAGGCTACCGCTCTTTGTAGGGGCTACACCTTC 1534

RESULT 6

US-09-188-930-259
; Sequence 259, Application US/09188930A

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew

Db	367	ACCACGATATGAACTGGAGAGATGGCATTTCACTGCAATGATATGATGATGCATCC	426
QY	758	TCCTGATGTTCTCTGCCAATATGATGATGTGTGAACACACCCGGCACATACCTTCTCTCT	817
Db	427	TCCTCGAGTTCCTCTGTCAACATAGTGTGAACACACCGGGCTCATACCTTCTCTCAT	486
QY	818	GCCCTCCAGGCTACATCCTGCTGTGATGACAACCGNAGCTGGCAAGATCATCAAGAAATGTG	877
Db	487	GCCTCCAGGCTWKTCTTGTGTGAAGATTAACCGAAGCTGCCAGATATCATATGATGTG	546
QY	878	AGCACAGAGAACCAACAGTGTCAACCTGCACGACGAGCTGTACATTTTACAAAGGGGGCTTCA	937
Db	547	AGCCACGGAAACCAACATGACACTCCCTCCGACACTTGTAAATTCGAAGGGGGCTTCA	606
QY	938	AATGATATGACCCCATCGCTGTGTGAGAGCCCTTATCTAGAGATCAGTATAACCGCTGTA	997
Db	607	AATGATATGACCCCATCGCTGTGTGAGAGCCCTTATCTCTGATTTGGGATTAACCGCTGTA	666
QY	998	TGTGTCTCTGTAGAAACCTGGCTGCAGAGACCAACCCCTTACATCTTTGACCGGGACA	1057
Db	667	TGTCCCTCGCTGAATAATCTGGCTGCAGAGACCAACCCCTTACATCTTTGTTCCGGGACA	726
QY	1058	TGGACGTGGTGTGAGGAGCGCTCCGTTCCGCTGTGACATTTTCCAATATGAACGCAGACC	1111
Db	727	TGGATGTGTGTATCAGAGACCTCTGTTCTGCTGACATTTTCCAGATGTGAAGAGACGCC	786
QY	1118	GCTACCTCTGGGGCTTATTCATTTTCCAGATCAATCTGGGATATGAGGGCAGAGAAATTT	1177
Db	787	GATACCTCTGGGGCTTATTCATTTTCCAGATCAATCTGGGATATGAGGGGTGAGAGTTCT	846
QY	1178	ACATGGGGCAAAAGGGGCCCATCGATCGTGCACCCGTGTATGACAGCCGCCCATCAAAAGGG	1233
Db	847	ACATGGGGCAAAAGGGGCTTATTCATGTGCACTCGTGTATGACAGCCGCCCATCAAAAGGG	906
QY	1238	CCCGGGAATTCACACTGTGACTTTGGAAATGATCACTGTCAACACTGTCTCAACTTCAGAG	1297
Db	907	CTCGGGACATTCACACTGTGACTTTGGAGATGATCACCGTCAACACTGTCTCAACTTCAGAG	966
QY	1298	GCACCTCGTATCGCACTGGGGATATATGTGTGCAATATACCATTCAT	1344
Db	967	GCACCTCGTATCGCACTGGGGATATATGTGTGCAATATACCATTCATTC	1013

```

? NAME: Brookes, A. Anders
? REGISTRATION NUMBER: 36.373
? REFERENCE/DOCKET NUMBER: PR258
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 301-309-8504
? TELEFAX: 301-309-8439
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1531 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 125..1453
?
US-08-833-963C-1

```

OY 1 ATGCCAGGAATAAAAAGATCTACCTGTACATTTCTGGCTCTCTGTCTTCCAGCCCT 60
DB 382 ATGCCAGGAATAAAAAGATCTACCTGTACATTTCTGGCTCTCTGTCTTCCAGCCCT 441
OY 61 GGGATATCAGACGACAGTGCAGCAATGGCTTTTGACATGATCGCAGTCAAGACAGTGT 120
DB 442 GGGATATCAGACGACAGTGCAGCAATGGCTTTTGACATGATCGCAGTCAAGACAGTGT 501
OY 121 TTGATATTTGATGAATGCCAGCAATCCCGAGAGCCCTCCGAGAGACATGATGTGT 180
DB 502 TTGATATTTGATGAATGCCAGCAATCCCGAGAGCCCTCCGAGAGACATGATGTGT 561
OY 181 AACCAAAATGGCGGATTTTATGATTCCTCCGAGCAAAACCTGTGTATCGAGGCCCTAC 240
DB 562 AACCAAAATGGCGGATTTTATGATTCCTCCGAGCAAAACCTGTGTATCGAGGCCCTAC 621
OY 241 TCGAACCCCTACTGACCCCTACTCAGGTCCGTACCCAGAGAGTGCCTCCACCTCTCA 300
b 622 TCGAACCCCTACTGACCCCTACTCAGGTCCGTACCCAGAGAGTGCCTCCACCTCTCA 681
OY 301 GCTCCAACTATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCAATGAT 360
DB 682 GCTCCAACTATCCCAAGATCTCCAGGCTCTTATATGCCCTTTGATACCAATGAT 741
OY 361 GAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCTCCAGAGTGCACCC 420
DB 742 GAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTCCTCCAGAGTGCACCC 801
OY 421 ACCAGATCTGATCAATCACTGAAGGCGGTACACCTGCTCTGACCCAGAGATTTGG 480
DB 802 ACCAGATCTGATCAATCACTGAAGGCGGTACACCTGCTCTGACCCAGAGATTTGG 861
OY 481 CTTCCTGAAGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 862 CTTCCTGAAGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 921
OY 541 TGTGCAATGTCTGATCTCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAG 600
DB 922 TGTGCAATGTCTGATCTCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAG 981
OY 601 GATGGAAGGTCTTGCCAGATGTGCAAGAGTGTGCCACCGAAGCCCTGCTGCAAAAC 660
DB 982 GATGGAAGGTCTTGCCAGATGTGCAAGAGTGTGCCACCGAAGCCCTGCTGCAAAAC 1041
OY 661 TGTGCAATGTCTGATCTCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAG 720
DB 1042 TGTGCAATGTCTGATCTCTTATCTTGTACATGCAACCTGGTTTACCCTCAATGAG 1101
OY 721 GATGCGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 1102 GATGCGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
OY 781 GATGTTGTGAACAGCCCGGACATATCTTGTCTGCTCCAGGCTACATCCTGCTG 840
DB 1162 GATGTTGTGAACAGCCCGGACATATCTTGTCTGCTCCAGGCTACATCCTGCTG 1221
OY 841 GATGCAACCGAGAGTGTGCAAGAGATGCAAGAGTGTGCAAGAGTGTGCAAGAGTGT 900
DB 1222 GATGCAACCGAGAGTGTGCAAGAGATGCAAGAGTGTGCAAGAGTGTGCAAGAGTGT 1281
OY 901 CTGCAAGACAGTGTCAATTTTACAAGGGGCTTCAATATGATGACCCCATCGGCT 960
DB 1282 CTGCAAGACAGTGTCAATTTTACAAGGGGCTTCAATATGATGACCCCATCGGCT 1341
OY 961 GAGGAGCCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1020
DB 1342 GAGGAGCCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1401
OY 1021 TGCAGAGACAGCCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1080
DB 1402 TGCAGAGACAGCCCTTATCTGAGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGGC 1461

OY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGCGCTATTACATT 1140
DB 1462 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGGGCGCTATTACATT 1521
OY 1141 TTCCAGATCAATTTGGGAAATGAGGGCGAGAAATTTTCAATGCGGCAAGGGGCCCTC 1200
DB 1522 TTCCAGATCAATTTGGGAAATGAGGGCGAGAAATTTTCAATGCGGCAAGGGGCCCTC 1581
OY 1201 AGTGCACCCCTGATGATGACAGCCCATCAAGAGGGCCCGGAAATCCAGCTGACATTG 1260
DB 1582 AGTGCACCCCTGATGATGACAGCCCATCAAGAGGGCCCGGAAATCCAGCTGACATTG 1641
OY 1261 GAATGATCACTGTCAACACCTGTATCACTTCAAGAGCAGCTCCGTATCCGACTGGG 1320
DB 1642 GAATGATCACTGTCAACACCTGTATCACTTCAAGAGCAGCTCCGTATCCGACTGGG 1701
OY 1321 ATATATGTCCGAGTACCCATTG 1344
DB 1702 ATATATGTCCGAGTACCCATTG 1725

RESULT 3
US-09-188-930-67
Sequence 67, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murlson, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ. ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO: 67
LENGTH: 1260
TYPE: DNA
ORGANISM: Rat
US-09-188-930-67
Query Match 62.63% Score 841.8; DB 3; Length 1260;
Best Local Similarity 89.43% Pred. No. 2.5e-259;
Matches 900; Conservative 5; Mismatches 102; Indels 0; Gaps 0;
OY 338 GCGGCTTTGGATCCAGATGATGAAAGCAACCAATGAGTGTGAGAGTGTGCA 397
DB 7 GTGCTTTGGATCCAGATGATGAAAGCAACCAATGAGTGTGAGAGTGTGCA 66
OY 398 CAGATTCACCAAGTGCACACCCAGATGATGATGATGATGATGATGATGATGATGATGAT 457
DB 67 CAGATTCACCAAGTGCACACCCAGATGATGATGATGATGATGATGATGATGATGATGAT 126
OY 458 GCTCTGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 517
DB 127 GCTCTGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
OY 518 GCTATGTTACTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCA 577
DB 187 GCTATGTTACTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCAAGAGTGTGCA 246
OY 578 ACCCTGTTTACCTCAATGAGATGAGAGTGTGCAAGAGTGTGCAAGAGTGTGCA 637
DB 247 ACCCTGTTTACCTCAATGAGATGAGAGTGTGCAAGAGTGTGCAAGAGTGTGCA 306
OY 638 CCGAGAACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTG 697
DB 307 CCGAGAACCCCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTGCAACCTGCTG 366
OY 698 ACCAGGATATGAACTTGAAGAGAGTGGCTTCAATGATGATGATGATGATGATGATGAT 757

Matches 1344: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGCAGGAATAAAGGATAGTCTGCTTACATTCCTGCTGCTTCCAAACCC 60
   |||||
Db 382 ATGCCAGGAATAAAGGATAGTCTGCTTACATTCCTGCTGCTTCCAAACCC 441
QY 61 GGAATGACACAGGACACATGACGAATGCTTGAACGTGATCGCCAGTACAGCAGTGT 120
   |||||
Db 442 GGAATGACACAGGACACATGACGAATGCTTGAACGTGATCGCCAGTACAGCAGTGT 501
QY 121 TTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
   |||||
Db 502 TTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
QY 181 AACCAAAATGCGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
   |||||
Db 562 AACCAAAATGCGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
QY 241 TCGAACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACT 300
   |||||
Db 622 TCGAACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACTGACCCCTACT 681
QY 301 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGATGATGATGATGATGATGATGATGAT 360
   |||||
Db 682 GCTCCAACTATCCACAGATCTCCAGGCTCTTATATGATGATGATGATGATGATGATGATGAT 741
QY 361 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
   |||||
Db 742 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 421 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
   |||||
Db 802 ACCCAGATCTGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 861
QY 481 CTTCGAAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
   |||||
Db 862 CTTCGAAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 921
QY 541 TGTGCGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
   |||||
Db 922 TGTGCGAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
QY 601 GATGGAAGTCTTGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
   |||||
Db 982 GATGGAAGTCTTGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
QY 661 TGGCTCAACACCTGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
   |||||
Db 1042 TGGCTCAACACCTGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101
QY 721 GATGCGGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
   |||||
Db 1102 GATGCGGTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
QY 781 GAGTGTGTAACAGCCCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
   |||||
Db 1162 GAGTGTGTAACAGCCCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221
QY 841 GATGACAACGGAAGCTGCAAGACATCAACGAATGTGACAGACAGACAGACAGACAGACAGAC 900
   |||||
Db 1222 GATGACAACGGAAGCTGCAAGACATCAACGAATGTGACAGACAGACAGACAGACAGACAGAC 1281
QY 901 CTGCAGACAGAGTCTACATTTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGAT 960
   |||||
Db 1282 CTGCAGACAGAGTCTACATTTTACAAAGGGGCTTCAATGATGATGATGATGATGATGATGATGAT 1341
QY 961 GAGAGGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
   |||||
Db 1342 GAGAGGCTTATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401
QY 1021 TGCAGAGACAGCCCTTACATCTTGTACCGGGACATGAGAGTGTGTACAGACCCCTCC 1080
   |||||
Db 1402 TGCAGAGACAGCCCTTACATCTTGTACCGGGACATGAGAGTGTGTACAGACCCCTCC 1461

```

```

QY 1081 GTTCCCGGTGACATCTTCCAAATGACAGCCAGACCCGCTACCTGGGCGCTTATCAT 1140
   |||||
Db 1462 GTTCCCGGTGACATCTTCCAAATGACAGCCAGACCCGCTACCTGGGCGCTTATCAT 1521
QY 1141 TTCCAGATCAAAATCTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1200
   |||||
Db 1522 TTCCAGATCAAAATCTGGGAATGAGGAGAGAAATTTTACATGCGGCAAAAGGGCCCATC 1581
QY 1201 AGTGCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
   |||||
Db 1582 AGTGCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
QY 1261 GAAATGATCACTGTCACACTGTCAATCACTGATGATGATGATGATGATGATGATGATGATGAT 1320
   |||||
Db 1642 GAAATGATCACTGTCACACTGTCAATCACTGATGATGATGATGATGATGATGATGATGATGAT 1701
QY 1321 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
   |||||
Db 1702 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1761

```

RESULT 2

US-09-212-168-2

Sequence 2, Application US/09212168

Patent No. 6303765

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Guebler, Karl J.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Inocyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/212,168

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/884,072

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2550 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: CORN001

CLONE: 45517

US-09-212-168-2

Query Match 100.0%; Score 1344; DB 4; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 62.0182 Seconds
(Without alignments)
6646.008 Million cell updates/sec

Sequence: 1 atgcacgaataaaagat.....atgtcgcagtaaccattc 1344

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents, NA: *

1: /cgn2-6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2-6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2-6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2-6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2-6/ptodata/1/lna/PCUS.COMB.seq:*
6: /cgn2-6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	2550	2	US-08-884-072-2
2	1344	100.0	2550	4	US-09-212-168-2
3	891.8	67.6	1260	3	US-09-188-930-67
4	371.2	27.6	1531	2	US-08-833-963C-1
5	371.2	27.6	2018	3	US-08-980-514-2
6	316.6	23.6	1018	3	US-09-188-930-259
7	241	17.9	2512	4	US-08-948-757-1
8	106.8	7.9	3373	2	US-08-897-443-2
9	89.4	6.7	401	4	US-09-643-597-278
10	80.2	6.0	220	4	US-09-404-879A-340
11	38.2	4.3	2461	1	US-08-282-141-1
12	38.2	4.0	3759	3	US-08-479-722B-3
13	52.4	3.9	5499	3	US-08-479-722B-1
14	52.4	3.9	5502	5	PCT-US95-02251-17
15	51.6	3.8	3753	5	PCT-US95-02251-2
16	51.6	3.8	4314	1	US-08-199-780-2
17	51.6	3.8	4314	2	US-08-316-650-2
18	51.6	3.8	5089	6	5177197-31
19	50.4	3.8	6412	4	US-08-652-877-17
20	50.4	3.8	6412	4	US-08-476-515A-17
21	50.4	3.8	14044	4	US-08-652-877-85
22	50.4	3.8	14044	4	US-08-652-877-89
23	50.4	3.8	14080	4	US-08-652-877-87
24	50.4	3.8	14083	4	US-08-476-515A-83
25	45.6	3.4	3460	2	US-08-652-877-83
26	45.6	3.4	3460	2	US-08-751-305-1
27	44.8	3.3	2181	6	5208144-36

28	43.4	3.2	3546	4	US-08-872-757-3	Sequence 3, Appl
29	41.4	3.1	2492	4	US-09-381-779-1	Sequence 1, Appl
30	39.8	3.0	3136	4	US-09-284-819-8	Sequence 8, Appl
31	38	2.8	6048	4	US-09-634-920-3	Sequence 3, Appl
32	37.8	2.8	1611	4	US-09-249-697A-2	Sequence 2, Appl
33	37.8	2.8	1611	4	US-09-363-316B-2	Sequence 2, Appl
34	37.8	2.8	2365	4	US-09-249-697A-5	Sequence 5, Appl
35	37.8	2.8	2365	4	US-09-249-697A-18	Sequence 18, Appl
36	37.8	2.8	2365	4	US-09-363-316B-5	Sequence 5, Appl
37	37.8	2.8	2365	4	US-09-363-316B-23	Sequence 23, Appl
38	37.2	2.8	2397	6	5258288-2	Patent No. 5258288
39	37	2.8	6344	4	US-08-843-417-1	Sequence 1, Appl
40	35.4	2.6	825	1	US-08-312-870-6	Sequence 6, Appl
41	35.4	2.6	1338	1	US-08-307-444A-8	Sequence 8, Appl
42	35.4	2.6	1338	1	US-08-587-389-8	Sequence 8, Appl
43	35.4	2.6	1368	1	US-08-307-444A-7	Sequence 7, Appl
44	35.4	2.6	1368	1	US-08-587-389-7	Sequence 7, Appl
45	35.4	2.6	1425	1	US-08-307-444A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-2
Sequence 2, Application US/08884072
Patent No. 5208144

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl J.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/884,072

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2550 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: CORNNOT01

CLONE: 45517

US-08-884-072-2

Query Match

Best Local Similarity 100.0%; Score 1344; DB 2; Length 2550;
Pred. No. 0;

Db	375	GCATCCAACTACCCCAAGATTTCAGAGCCCTTGTGTGGGCTTTGGGTATTCAGATGAT	4334
Qy	361	GAAAGCAACCAATGTGTGGATGTGGACAGAGTGTGCAACAGATTTCCACAGTGCACACCC	420
Db	435	GAAGGCACACAGTGTGGATGTGGACAGAGTGTGCAACACAGCTCACACCCAGTGCACACCT	494
Qy	421	ACCCAGATCTGCATCAATACATGAAGCGGGGTACACCGCTCCAGACCGGAGGATATTTGG	480
Db	495	ACCCAGATCTGTATCAACACTGGAAGAGGTATACACTGCTCTCTGACACCGATGGGTACTGG	554
Qy	481	CTTCTGGAAGCCAGTGTGTAGACATATGATGAATGTGCTATGTTACTGTACTGCAGAGCTC	540
Db	555	CTTCTGGAAGGGGAGTGGCTAGATATTTGATGAATGTGCTATGTTACTGTACTGCAGAGCTC	614
Qy	541	TGTGGCAATGTTCCTGGATTCCTATTTCTTGTGACATGCAACCTGTGTTTACCTTCATAGAG	600
Db	615	TGTGCAAAATGTTCAGAGATCTATTTCTGTACATGCAACCTGTGTTTACCTTCACACGAGAC	674
Qy	601	GATGGAAGGTCTTGTCCAAAGATGGAACGAGTGTGCCACGAGAAACCCCTGCTGCACACAC	660
Db	675	GATGGAAGGTCTTGTCCAAAGATGGAACGAGTGTGCCAAAGTGAAGATCCCTGTGTTCAGACC	734
Qy	661	TGCGTCAACACTACAGGCTCTTTCACTGTCCGCTGTGACCCAGAGATATGAATTTGAGAA	720
Db	735	TGTGTCAACACTATGTGGCTTTTCATCTCCGCTGTGACCCAGAGATATGAATTTGAGAA	794
Qy	721	GATGGCTTCATTGTGCAATATGACGAGTGAAGTTCCTAGATGTTCCCTGTGCCAACAT	780
Db	795	GATGGCTTCATTGTGCAATATGACGAGTGAAGTTCCTAGATGTTCCCTGTGTCAACAC	854
Qy	781	GAGTGTGTGAACCAAGCCCGGCAATATCTTGTCTCTGCTCCAGCTACATCTCTGCTG	840
Db	855	GAGTGTGTGAACCAAGCCCGGCTCATCTTCTGCTCTGCTCCAGCTACATCTCTGCTG	914
Qy	841	GATGCAACCGGAAGCTGGCCAAAGATCAACAGAAATGTGACACAGGAACCAACAGCTGCAAC	900
Db	915	GATGCAACCGGAAGCTGGCCAGATATCAATGAATGTGACACCGGAACCAACAGCTGTAC	974
Qy	901	CTGCAAGCAGACGTGTCTACAAATTTTACAAAGGGGGCTTCAATGTGATGCAAGCCCATCCGCTGT	960
Db	975	TCACGTGCAAGACTTGTGTACAAATCTACAAAGGGGGCTTCAATGTGATGCAAGCTCATCAGCTGT	103
Qy	961	GAGGAGCCCTTATCTGAAGATTCAGTGTAAACCGCTGTATGTGTCTGTGTGAGAAACCTTGGC	102
Db	1035	GAGGAGCCCTTATCTGTCTGTATGTGTGTAACACCCCTGTATGTGTCTGTGTGAGCAACACAGC	109
Qy	1021	TGCAAGAGCAGACCCCTTACATCTTGTATCCGGACATGAGAGTGTGTGACAGAGCTGC	108
Db	1095	TGCAAGAGCAGACCCCTTACATCTGTATGTGTGAGAGATGTGTGTGTGTGTGTGTGTGTGTGT	115
Qy	1081	GTTTCCCGCTGACATCTTCCAAATGCAAGCCAGACCGCTACCTGTGGGGCTTATTACAT	1144
Db	1155	GTTTCTGTGTGACATCTTCCAGATGCAACCAACACCCGATACCTGTGGTCTTATTACAT	121
Qy	1141	TTTCCAGATCAAAATCTGGGAATGTAGGGCAGAGAAATTTTACATGTGGGCAAAACGGGCCATC	120
Db	1215	TTTCCAGATCAAAATCTGGGACAGAGGGTGTGAGAGTCTATATGTGGGCAAAACAGGGGCTATC	127
Qy	1201	AGTGCACCCCTGTGTATGACACGCCCCATCAAAAGGGGCCCGGGAATCTCAGTGTGACTTG	126
Db	1275	AGTGCACCCCTGTGTATGACACGCCCCATCAAAAGGGGCTCGGGAATCTCAGTGTGACTTG	133
Qy	1261	GAAATGATCACTGTCAACACTGTCTATCAACTTCAGAGGAGCTCCGTATCCGACTGTGGG	1320
Db	1335	GAGATGATCACTGTCAACACTGTCTATCAACTTCAGAGGAGCTCCGTATCCGACTGTGGG	139
Qy	1321	ATATATGTGTGCAAGTACCACTTC	1344
Db	1395	ATATATGTGTGCAAGTATCCGTT	1418

ID	AAZ40029	standard; DNA; 2233 BP.
XX	AAZ40029;	
XX	15-FEB-2000	(first entry)
XX	Full length mouse A55 protein coding sequence.	
XX	A55 protein; mouse; smooth muscle proliferation; tissue generation;	
XX	vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;	
XX	vascular endothelial thickening; haematopoietic cell-regulator; cytokine;	
XX	percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;	
XX	actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;	
XX	tumour metastasis inhibitor; ss.	
XX	Mus musculus.	
XX	MO955864-A1.	
XX	04-NOV-1999.	
XX	28-APR-1999;	99WO-JP02284.
XX	28-APR-1998;	98JP-0119731.
XX	(ONOX) ONO PHARM CO LTD.	
XX	Honjo T, Tashiro K, Nakamura T;	
XX	WPI: 2000-038647/03.	
XX	P-PSDB; AAY54990.	
XX	Novel human polypeptides for treatment of, e.g. arteriosclerosis and	
XX	myoma -	
XX	Example 4; Page 52-55; 87pp; Japanese.	
XX	This sequence encodes the mouse A55 protein. The invention relates to the	
XX	human A55 protein. The protein can be used for the treatment of diseases	
XX	due to abnormal proliferation of smooth muscle. The polypeptides can be	
XX	used according their inhibition of the proliferation of vascular smooth	
XX	muscle cells, particularly in treating arteriosclerosis or re-narrowing	
XX	by vascular endothelial thickening after percutaneous transluminal	
XX	coronary angioplasty (PTCA), or myoma, haematopoietic cell-regulatory	
XX	activity, cytokine activity, tissue generation/repairation activity,	
XX	actin/inhibin activity, taxis and chemotaxis activity, blood	
XX	coagulation/thrombotic activity, receptor/ligand activity,	
XX	cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as	
XX	a nutrient.	
XX	Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other:	
XX	Query Match	82.9%; Score 1113.6; DB 21; Length 2233;
XX	Best Local Similarity	89.3%; Pred. No. 0;
XX	Matches 1200; Conservative	0; Mismatches 144; Indels 0; Gaps 0
YY	1 ATGCCAGGAATAAAGAGTACTCAGCTTACCATCTTGGGCTCTGTCTTCCAAGCGCT	60
YY		
YY	DB 75 ATGCCAGGAATAAAGAGTACTCAGCTTACCATCTTGGGCTCTGTCTTCCAACATCT	134
YY	61 GGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATMCGGCAGTCAGGACAGTGT	120
YY		
YY	DB 135 GGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATMCGGCAGTCAGGACAGTGT	194
YY	121 TTGATATTGTGATGATGCCGGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGT	180
YY		
YY	DB 195 CTGATATTGTGATGATGCCGGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTGC	254
YY	181 AACCAAAATGCGGGGTATTTATGATATCCCGGACAAACCTGTGTATCGAGGGCGCTTAC	240
YY		
YY	DB 255 AACCAAAATGCGGGGTATTTGTGATATCCCGGACAAACCTGTGTATCGAGGGCGCTTAC	314

```

QY 421 ACCGAGATCTGACATCAATCTGAAGCGGGTACACCTGCTCTGGACCGAGCATATATGG 480
D 421 ACCGAGATCTGATCAACACTGAAGGAGTTACACTGCTCTGGACCGAGTGGTACTGG 480
QY 481 CTCTGAGAGCGCCAGTCTTGAACATTCATGATTCCTGCTATGTTACTGCCAGCAGCTC 540
D 481 CTCTGAGAGCGCCAGTCTTGAACATTCATGATTCCTGCTATGTTACTGCCAGCAGCTC 540
QY 541 TGTGCAATGTTCTGATCTTCTGATGATGAGAGAGTGCACCTGCTTTCACCTCATATGAG 600
D 541 TGTGCAATGTTCTGATCTTCTGATGATGAGAGTGCACCTGCTTTCACCTCATATGAG 600
QY 601 GATGAGAGTCTTCCCAAGATGTGACAGAGTGCACCGAGACCCCTGCGTGAACCC 660
D 601 GATGAGAGTCTTCCCAAGATGTGACAGAGTGCACCGAGACCCCTGCGTGAACCC 660
QY 661 TGCCTCAACACCTTACCGCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
D 661 TGCCTCAACACCTTACCGCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAA 720
QY 721 GATGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 780
D 721 GATGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 780
QY 781 GAGTGTGAAACGACCGCGGACATCTTCTGCTCTGCGCTCCAGAGCTACATCTGCTG 840
D 781 GAGTGTGAAACGACCGCGGACATCTTCTGCTCTGCGCTCCAGAGCTACATCTGCTG 840
QY 841 GATGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 900
D 841 GATGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 900
QY 901 CTGAGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 960
D 901 CTGAGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 960
QY 961 GAGGAGCCTTATCTGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1020
D 961 GAGGAGCCTTATCTGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1020
QY 1021 TGCAGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1080
D 1021 TGCAGAGAGTCTTACGATGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1080
QY 1081 GTTCCGCTGACATCTTCCAAATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1140
D 1081 GTTCCGCTGACATCTTCCAAATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1140
QY 1141 TTTCCAGATCAATCTGAGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1200
D 1141 TTTCCAGATCAATCTGAGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGAT 1200
QY 1201 AGTGGCCACCTGAGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGATGAT 1260
D 1201 AGTGGCCACCTGAGATGAGAGTGCACCTGCTTTCATGATGATGATGATGATGATGAT 1260
QY 1261 GAATATGATCTGATCAACACTGCTATCACTTCAAGAGGAGTCCGTGATCCGATGCGG 1320
D 1261 GAATATGATCTGATCAACACTGCTATCACTTCAAGAGGAGTCCGTGATCCGATGCGG 1320
QY 1321 ATATATGATCTGATCAACACTGCTATCACTTCAAGAGGAGTCCGTGATCCGATGCGG 1380
D 1321 ATATATGATCTGATCAACACTGCTATCACTTCAAGAGGAGTCCGTGATCCGATGCGG 1380

```

RESULT 14
AAZ39384

ID AAZ39384 standard; DNA: 2233 BP.

AC AAZ39384;

XX 22-FEB-2000 (first entry)

```

XX DE Smooth muscle proliferation modulating protein encoding DNA.
XX KW Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
XX KW endothelial thickening; percutaneous transluminal coronary angioplasty;
XX KW myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
XX KW actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
XX KW metastasis; nutrient; ss.
XX OS Mus musculus.
XX PN M0955863-A1.
XX PD 04-NOV-1999.
XX PE 28-APR-1999; 99WO-JP02283.
XX PR 28-APR-1998; 98JP-0119731.
XX PA (ONOX ) ONO PHARM CO LTD.
PI Honjo T, Tashiro K, Nakamura T;
XX WPI: 2000-038646/03.
DR P-PSDB: AAY56750.
PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
PS Claim 5; Page 47-50; 70pp; Japanese.
XX The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/reproduction
CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents a DNA encoding the protein of the
CC invention which can be used for modulating smooth muscle cell
CC proliferation.
SQ Sequence 2233 BP; 554 A; 582 C; 510 G; 586 T; 1 other:
Query Match 82.9%; Score 1113.6; DB 21; Length 2233;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 1 ATGCCAGAAATTAAGATATCTACCTGTTACATTTGGCTCTGCTTCCAAACCTT 60
D 1 ATGCCAGAAATTAAGATATCTACCTGTTACATTTGGCTCTGCTTCCAAACCTT 60
QY 75 ATGCCAGAAATTAAGATATCTACCTGTTACATTTGGCTCTGCTTCCAAACCTT 134
D 75 ATGCCAGAAATTAAGATATCTACCTGTTACATTTGGCTCTGCTTCCAAACCTT 134
QY 61 GGAATGACAGCGGACAGTGCACGAAATGCTTGAACCTGATGCGCAGTCAAGAGTGT 120
D 61 GGAATGACAGCGGACAGTGCACGAAATGCTTGAACCTGATGCGCAGTCAAGAGTGT 120
QY 135 GGGATGACAGCGGACAGTGCACGAAATGCTTGAACCTGATGCGCAGTCAAGAGTGT 194
D 135 GGGATGACAGCGGACAGTGCACGAAATGCTTGAACCTGATGCGCAGTCAAGAGTGT 194
QY 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
D 121 TTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 181 AACCAAAATGGGGGATTTTATGATTTCCCGGACAAACCTGATGATGATGATGATGAT 240
D 181 AACCAAAATGGGGGATTTTATGATTTCCCGGACAAACCTGATGATGATGATGATGAT 240
QY 241 TCGAAGCCCTTACGACCCCTTACGACGATGCGTCCGAGAGTCCGAGAGTCCGAGAGTCC 300
D 241 TCGAAGCCCTTACGACCCCTTACGACGATGCGTCCGAGAGTCCGAGAGTCCGAGAGTCC 300
QY 315 TCAAAATCCCTACTTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 374
D 315 TCAAAATCCCTACTTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 374
QY 301 GCTCCAAACTATCCAGATCTCCAGAGCTCTTATATGCGGCTTGGATACGATGAT 360

```

```

|||||
362 CAGGCTTAGACATGATGATGATGCGATGTTACTGCGACACTGTCGGAATGTT 421
OY 553 CCGGATTCCTATTTTGTACATGCAACCTGTTTACCTCCTAATGAGATGGAAGTCT 612
DB 422 CCGGATTCCTATTTTGTACATGCAACCTGTTTACCTCCTAATGAGATGGAAGTCT 481
OY 613 TGGCAAGATGTAAGAGAGTGTGGCAGGAGAACCCCTGCGCAAACTGCGCTCAACAC 672
DB 482 TGGCAAGATGTAAGAGAGTGTGGCAGGAGAACCCCTGCGCAAACTGCGCTCAACAC 541
OY 673 TACGGCTCTTTCATCTGCGCTGAGCCAGGATATGAACTTGAGGAAGATGCGCTTCAT 732
DB 542 TACGGCTCTTTCATCTGCGCTGAGCCAGGATATGAACTTGAGGAAGATGCGCTTCAT 601
OY 733 TGCAGTATATGAGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGATGTTGTAAC 792
DB 602 TGCAGTATATGAGACGAGTGCAGCTTCTCTGAGTTCTCTGCCAATGATGTTGTAAC 661
Y 793 CAGCCGGACATATCTGCTGCTGCTGCGCTCAGGCTCATCTGCTGGATGACAACGA 852
DB 662 CAGCCGGACATATCTGCTGCTGCTGCGCTCAGGCTCATCTGCTGGATGACAACGA 721
OY 853 AGCTGCCAAGACATCAACGAATGTGACACAGAACCAAGTGCACCTGCAGACAGAG 912
DB 722 AGCTGCCAAGACATCAACGAATGTGACACAGAACCAAGTGCACCTGCAGACAGAG 781
OY 913 TGTACAAATTTACAGAGGGGCTTAAATGCATGCAACCCCTGCTGAGAGAGCTTAT 972
DB 782 TGTACAAATTTACAGAGGGGCTTAAATGCATGCAACCCCTGCTGAGAGAGCTTAT 841
OY 973 CTGAGATCAGTATACCGCTGTATGTGCTGCTGAGAACCCCTGCTGAGAGAGAG 1032
DB 842 CTGAGATCAGTATACCGCTGTATGTGCTGCTGAGAACCCCTGCTGAGAGAGAG 901
OY 1033 CCCTTTACCATCTTGTACCGGGAGACATGAGCTGTGTCAGAGAGCTCCGTTCCGCTGAC 1092
DB 902 CCCTTTACCATCTTGTACCGGGAGACATGAGCTGTGTCAGAGAGCTCCGTTCCGCTGAC 961
OY 1093 ATCTTCCAATGCAAGCCAGACAGCCGCTACCCCTGAGGCTTATACATTTTCCAGATCAA 1152
DB 962 ATCTTCCAATGCAAGCCAGACAGCCGCTACCCCTGAGGCTTATACATTTTCCAGATCAA 1021
OY 1153 TCTGGAAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGCCCATCAGTCCACCCGTG 1212
DB 1022 TCTGGAAATGAGGGCAGAGATTTTACATGCGGCAAAAGGGCCCATCAGTCCACCCGTG 1081
OY 1213 GTGATGACACGCCCCCATCAAAAGGGCCCGGAAATCCAGCTGACCTGGAATGATCACT 1272
DB 1082 GTGATGACACGCCCCCATCAAAAGGGCCCGGAAATCCAGCTGACCTGGAATGATCACT 1141
OY 1273 GTCAACATCTCATCAACTTCAGAGGAGAGCTCCCTGATCCGATTCGGATATATGTGTCG 1332
DB 1142 GTCAACATCTCATCAACTTCAGAGGAGAGCTCCCTGATCCGATTCGGATATATGTGTCG 1201
OY 1333 CAGTACCATTC 1344
DB 1202 CAGTACCATTC 1213

```

```

KW actin activity; Chemotaxis; blood coagulation; thrombotic; tumor; mouse;
KM metastasis; nutrient; ss.
XX Mus musculus.
XX WO955863-A1.
XX 04-NOV-1999.
XX 28-APR-1999; 99WO-JP02283.
XX 28-APR-1998; 98JP-0119731.
XX (ONOV ) ONO PHARM CO LTD.
XX Honjo T, Tashiro K, Nakamura T;
XX WPI; 2000-038646/03.
XX P-PSDB; AAY56750.
XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
PT Claim 4; Page 45-46; 70pp; Japanese.
XX PS The invention provides mouse polypeptides for treatment of diseases due
XX CC to abnormal proliferation of smooth muscle. The polypeptides can be
XX CC produced by standard recombinant methodology. The polypeptides can be
XX CC used according to their inhibition of the proliferation of vascular
XX CC smooth muscle cells, particularly in treating arteriosclerosis or re-
XX CC narrowing by vascular endothelial thickening after percutaneous
XX CC transluminal coronary angioplasty (PTCA), or myoma, hemipoletic cell-
XX CC regulatory activity, cytokine activity, tissue generation/repair
XX CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
XX CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
XX CC tumor metastasis inhibiting activity; tumor inhibition, and as nutrient.
XX CC The present sequence represents the coding sequence of one protein of the
XX CC invention which can be used for modulating smooth muscle cell
XX CC proliferation.
SQ Sequence 1344 BP; 325 A; 370 C; 330 G; 319 T; 0 other;
Query Match 82.9%; Score 113.6; DB 21; Length 1344;
Best Local Similarity 89.3%; Pred. NO. 0;
Matches 1200; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
OY 1 ATGCCAGGAATTAAGAGATACATCTTACATCTTGGCTCTGTTCCAGGCTT 60
DB 1 ATGCCAGGAATTAAGAGATACATCTTACATCTTGGCTCTGTTCCAGGCTT 60
OY 61 GGGAAATGACAGGACGATGCGATGCTTTGACCTGATGCCAATCAGAGAGTGT 120
DB 61 GGGAAATGACAGGACGATGCGATGCTTTGACCTGATGCCAATCAGAGAGTGT 120
OY 121 TTAGATATTTGATGATGCGGACATCCCGAGGCTTCGAGGAGAGATGATGTGT 180
DB 121 CTGATATTTGATGATGCGGACATCCCGAGGCTTCGAGGAGAGATGATGTGT 180
OY 181 AACCAAAATGCGGGATTTATGATTCCTCCCGGACAAACCTGTGTATCGAGGCGCTTAC 240
DB 181 AACCAAAATGCGGGATTTATGATTCCTCCCGGACAAACCTGTGTATCGAGGCGCTTAC 240
OY 241 TCGAACCCCTACTGACAGCCCTTACTAGGTCCGTAACCGAGAGCTGCCCACTCTTCA 300
DB 241 TCGAACCCCTACTGACAGCCCTTACTAGGTCCGTAACCGAGAGCTGCCCACTCTTCA 300
OY 301 GGTCCAAATGATGCGGATTCAGAGGCTTCTTATGCGGCTTGTATACAGATGAT 360
DB 301 GGTCCAAATGATGCGGATTCAGAGGCTTCTTATGCGGCTTGTATACAGATGAT 360
OY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCACCC 420
DB 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAACAGATTTCCACAGTGCACCC 420

```

RESULT 13
 AA239383
 ID AA239383 standard; DNA; 1344 BP.
 XX
 AC AA239383;
 XX
 DE 22-FEB-2000 (first entry)
 XX
 DE Smooth muscle proliferation modulating protein coding sequence.
 XX
 KW Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 KW endothelial thickening; percutaneous transluminal coronary angioplasty;
 KW myoma; hemipoletic cell-regulation; cytokine; tissue generation; taxis;

PR 30-MAR-2000; 2000MO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194447P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000MO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000MO-US23328.
 PR 08-NOV-2000; 2000MO-US30952.
 PR 01-DEC-2000; 2000MO-US32678.
 PR 20-DEC-2000; 2000MO-US34956.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL:
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI: 2001-602746/68.
 P-PSDB: AAU29227.

Novel nucleic acids encoding pro polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 2: Fig 407; 774pp; English.

Sequences MAS45925-MAS46231 represent DNA molecules encoding and PCR primers for pro polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a pro polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The pro proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the pro nucleic acids can be used for genetic analysis of individuals with genetic disorders.

Sequence 2609 BP: 622 A: 718 C: 622 G: 647 T: 0 other:

Query Match 99.98; Score 1342.4; DB 22; Length 2609;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCCAGAAATTAAGACTACTGCTTACCATCTGGCTCTGTCTTCAGAGCCCT 60
 |||||||
 DB 450 ATGCCAGAAATTAAGACTACTGCTTACCATCTGGCTCTGTCTTCAGAGCCCT 509
 OY 61 GGGAAATGACAGCAGCAGTGCAGTGCCTTGGACCTGGATGCCAGTGCAGACAGTGT 120
 |||||||
 DB 510 GGGAAATGACAGCAGCAGTGCAGTGCCTTGGACCTGGATGCCAGTGCAGACAGTGT 569
 OY 121 TTAGATATTGATGATGCGGACATCCCGGAGGCGCTGCCGAGAGACATGATGTGTGT 180
 |||||||

DB 570 TTAGATATTGATGATGCGGACATCCCGGAGGCGCTGCCGAGAGACATGATGTGTGT 629
 OY 181 AACCAAAATGGGGGATTTATGATTCCTCCCGGAGAACCCGTGTATGAGGGCCCTAC 240
 |||||||
 DB 630 AACCAAAATGGGGGATTTATGATTCCTCCCGGAGAACCCGTGTATGAGGGCCCTAC 689
 OY 241 TCGAACCCCTACTGACACCCCTACTCAGGTCCGTACCCAGCAGCTCCGACCTCTCA 300
 |||||||
 DB 690 TCGAACCCCTACTGACACCCCTACTCAGGTCCGTACCCAGCAGCTCCGACCTCTCA 749
 OY 301 GCTCCAAATCTATCCAGATCTCCAGGCTCTTATATGCTGCTTGGATACAGATGAT 360
 |||||||
 DB 750 GCTCCAAATCTATCCAGATCTCCAGGCTCTTATATGCTGCTTGGATACAGATGAT 809
 OY 361 GAAGCAACCAATGTGTGATGTGAGCGATGTGCAACAGATTCCTCCAGATCAACCCC 420
 |||||||
 DB 810 GAAGCAACCAATGTGTGATGTGAGCGATGTGCAACAGATTCCTCCAGATCAACCCC 869
 OY 421 ACCGATCTGCATCATCTGATGAGCGGATACACTGCTGCTGCAACCGAGATATGG 480
 |||||||
 DB 870 ACCGATCTGCATCATCTGATGAGCGGATACACTGCTGCTGCAACCGAGATATGG 929
 OY 481 CTTCTGGAAGGCGAGTCTTACATTTGATGATGCTGTATGCTTACTGCTCCAGAGCTC 540
 |||||||
 DB 930 CTTCTGGAAGGCGAGTCTTACATTTGATGATGCTGTATGCTTACTGCTCCAGAGCTC 989
 OY 541 TGTGGATGTTCTCGATGATCTTATCTTGTATGATGACCCCTGCTTACCCCTCATAG 600
 |||||||
 DB 990 TGTGGATGTTCTCGATGATCTTATCTTGTATGATGACCCCTGCTTACCCCTCATAG 1049
 OY 601 GATGAGAGGTCTTGGCAAGATGTGAACGAGTGTGCACGAGAACCCCTGCTGCAACCC 660
 |||||||
 DB 1050 GATGAGAGGTCTTGGCAAGATGTGAACGAGTGTGCACGAGAACCCCTGCTGCAACCC 1109
 OY 661 TCGCTCAACACCTACAGGCTCTTACATCTCCGCTGTGACCCAGATATGATGAGGA 720
 |||||||
 DB 1110 TCGCTCAACACCTACAGGCTCTTACATCTCCGCTGTGACCCAGATATGATGAGGA 1169
 OY 721 GATGAGGTCTTACATGATGATGAGAGGATGAGTGTGCTGCTGCTGCTGCTGCAACAT 780
 |||||||
 DB 1170 GATGAGGTCTTACATGATGATGAGAGGATGAGTGTGCTGCTGCTGCTGCTGCAACAT 1229
 OY 781 GAGTGTGAAACACCGCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 |||||||
 DB 1230 GAGTGTGAAACACCGCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1289
 OY 841 GATGACACCGAAGTGTGCAACATCAACGATGAGTGTGAGACAGCAACAGAGTGTGAC 900
 |||||||
 DB 1290 GATGACACCGAAGTGTGCAACATCAACGATGAGTGTGAGACAGCAACAGAGTGTGAC 1349
 OY 901 CTGACGACAGCTGTCTACATTTTACAAAGGGGCTTCAAAATGATGACCCATCCGCTGT 960
 |||||||
 DB 1350 CTGACGACAGCTGTCTACATTTTACAAAGGGGCTTCAAAATGATGACCCATCCGCTGT 1409
 OY 961 GAGAGCTTATCTGAGATCATGATTAACCCCTGTATGTCTGCTGCTGAGAACCTGGC 1020
 |||||||
 DB 1410 GAGAGCTTATCTGAGATCATGATTAACCCCTGTATGTCTGCTGCTGAGAACCTGGC 1469
 OY 1021 TGCAGAGACAGCCCTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1080
 |||||||
 DB 1470 TGCAGAGACAGCCCTTACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1529
 OY 1081 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTTATACAT 1140
 |||||||
 DB 1530 GTTCCGCTGACATCTTCCAAATGCAAGCCAGACCCGCTACCTGCGGCTTATACAT 1589
 OY 1141 TTCCAGATCAAAATCTGGGATGAGGCGAGAGATTTTACATGCGGCAACGGCCCATC 1200
 |||||||
 DB 1590 TTCCAGATCAAAATCTGGGATGAGGCGAGAGATTTTACATGCGGCAACGGCCCATC 1649
 OY 1201 AGTGCACCTGTGATGATGACAGCCCATCAAGGCGCCGCGGGAATTCAGTGGACTTG 1260
 |||||||
 DB 1650 AGTGCACCTGTGATGATGACAGCCCATCAAGGCGCCGCGGGAATTCAGTGGACTTG 1709

CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
 CC hypersensitivity pneumonitis, and transplantation associated diseases
 CC (graft rejection, and graft-versus-host-disease). (I), its (ant)agonists
 CC or fragment can also be used as an adjuvant in treatment of tumors.
 CC Antibodies against (I) can also be used for diagnosing such diseases.
 CC This sequence encodes a human EGF-like homologue (Pro217) encoded by cDNA
 CC clone DMS2279 which is described in the invention.

XX Sequence 2509 BP; 591 A; 695 C; 590 G; 633 T; 0 other;

Query Match 99.9%; Score 1342.4; DB 20; Length 2509;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1343; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

  1 ATGCCAGGAATMAAAGATGATCTCACTTTACATTTCTGGCTCTGTCTTCCAAAGCCCT 60
  450 ATGCCAGGAATMAAAGATGATCTCACTTTACATTTCTGGCTCTGTCTTCCAAAGCCCT 509
  61 GGGATGCAAGGACAGAGTGCAGATGGCTTTGACCTGGATGGCCAGTGAAGACAGTGT 120
  510 GGGATGCAAGGACAGAGTGCAGATGGCTTTGACCTGGATGGCCAGTGAAGACAGTGT 569
  121 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
  570 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
  181 AACCAAAATGCGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 240
  630 AACCAAAATGCGGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 689
  241 TCGAACCCTTACTGAGCCCTTACTGAGGCTGATGATGATGATGATGATGATGATGATGAT 300
  690 TCGAACCCTTACTGAGCCCTTACTGAGGCTGATGATGATGATGATGATGATGATGATGAT 749
  301 GCTCCAAATCTATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
  750 GCTCCAAATCTATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809
  361 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
  810 GAAAGCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 869
  421 ACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
  870 ACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 929
  481 CTTCTGGAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
  930 CTTCTGGAAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 989
  541 TGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
  990 TGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1049
  601 GATGAAGGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
  1050 GATGAAGGCTTGTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1109
  661 TGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
  1110 TGTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1169
  721 GATGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
  1170 GATGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
  781 GACTGTGTGAACGAGCCGACATGATGATGATGATGATGATGATGATGATGATGATGAT 840
  1230 GACTGTGTGAACGAGCCGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
  841 GATGACAACGAGGCTGCAAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 900
  
```

```

  Db 1290 GATGACAACGAGAGCTGCCAAGACATCAAGATGATGACACAGAAACCAACAGTCAC 1349
  QY 901 CTGACAGACAGCTGCTATCAATTTTAAAGGGGCTTAAATGATGATGATGATGATGATGAT 960
  Db 1350 CTGACAGACAGCTGCTATCAATTTTAAAGGGGCTTAAATGATGATGATGATGATGATGAT 1409
  QY 961 GAGGAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
  Db 1410 GAGGAGCCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1469
  QY 1021 TGCAGAGACAGCCCTTTACATCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1080
  Db 1470 TGCAGAGACAGCCCTTTACATCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1529
  QY 1081 GTTCCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
  Db 1530 GTTCCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1589
  QY 1141 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGGGCCCATC 1200
  Db 1590 TTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGCGGCAACGGGCCCATC 1649
  QY 1201 AGTGCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
  Db 1650 AGTGCACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1709
  QY 1261 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
  Db 1710 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769
  QY 1321 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
  Db 1770 ATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1793
  
```

RESULT 10
 AAS46128
 ID AAS46128 standard; cDNA; 2609 BP.
 XX
 AC AAS46128;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human DNA encoding PRO polypeptide sequence #204.
 XX
 KW PRO polypeptide: mammal; tumour; cancer; human; cattle; horse; sheep; ss;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 PN WO200168848-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 28-FEB-2001; 2001WO-US06520.
 XX
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 28-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 29-MAR-2000; 2000US-193053P.

QY 361 GAAAGCAACCAATGTGTGATGTGACGAGTGTGCAACAGATTCCCAACAGTCCCAACCCC 420
DB 742 GAAAGCAACCAATGTGTGATGTGACGAGTGTGCAACAGATTCCCAACAGTCCCAACCCC 801
QY 421 ACCCAGATCTGCATCAATCTGAGAGCGGGTACACCTGCTCTGACCGACGAGATTTGG 480
DB 802 ACCCAGATCTGCATCAATCTGAGAGCGGGTACACCTGCTCTGACCGACGAGATTTGG 861
QY 481 CTTCTGGAAGCCAGTGTGATGACATGTGAAATGCGCTATGGTGTACTGCGACGAGCTC 540
DB 862 CTTCTGGAAGCCAGTGTGATGACATGTGAAATGCGCTATGGTGTACTGCGACGAGCTC 921
QY 541 TGTGCGAATGTTCTGTGATCTTATCTGTACATGCAACCCGTGTTTACCCCTCAATGAG 600
DB 922 TGTGCGAATGTTCTGTGATCTTATCTGTACATGCAACCCGTGTTTACCCCTCAATGAG 981
QY 601 GATGGAAGTCTTGGCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTCAAAACC 660
DB 982 GATGGAAGTCTTGGCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTCAAAACC 1041
QY 661 TGGCTCAACACCTACGGCTTTTCACTGCGCTGTGACCCGAGATATGAACTTGAAGAA 720
DB 1042 TGGCTCAACACCTACGGCTTTTCACTGCGCTGTGACCCGAGATATGAACTTGAAGAA 1101
QY 721 GATGCGCTTCAATTCAGTATATGAGACGAGTGTGACGCTTCTGAGTTCTCTGCAACAT 780
DB 1102 GATGCGCTTCAATTCAGTATATGAGACGAGTGTGACGCTTCTGAGTTCTCTGCAACAT 1161
QY 781 GAGTGTGTGAACCAACCCGCGACATATCTTCTCTCTCTCCCTCCAGAGTACATCTGCTG 840
DB 1162 GAGTGTGTGAACCAACCCGCGACATATCTTCTCTCTCTCCCTCCAGAGTACATCTGCTG 1221
QY 841 GATGACAACCGAAGTGTGCCAAGACATCAACGAATGTGACAGAGAACCAACAGCTGCAAC 900
DB 1222 GATGACAACCGAAGTGTGCCAAGACATCAACGAATGTGACAGAGAACCAACAGCTGCAAC 1281
QY 901 CTGACAGACGAGTGTGTACATTTTACAAAGGGGGCTTCAAAATGATGACACCCATCCGCTGT 960
DB 1282 CTGACAGACGAGTGTGTACATTTTACAAAGGGGGCTTCAAAATGATGACACCCATCCGCTGT 1341
QY 961 GAGAGAGCTTATCTGAGAGATCAAGTAAACCCCTGTATGTCTGCTGAGAAACCTGTGC 1020
DB 1342 GAGAGAGCTTATCTGAGAGATCAAGTAAACCCCTGTATGTCTGCTGAGAAACCTGTGC 1401
QY 1021 TGCAGAGACGAGCCCTTACCACTTTGTACCGGAGCATGAGGTGTGACAGAGCTGC 1080
DB 1402 TGCAGAGACGAGCCCTTACCACTTTGTACCGGAGCATGAGGTGTGACAGAGCTGC 1461
QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCCCGCTTACCTGAGGCTTATTTACAT 1140
DB 1462 GTTCCCGCTGACATCTTCCAAATGCAAGCCAGCCCGCTTACCTGAGGCTTATTTACAT 1521
QY 1141 TTCCAGATCAATCTGGGAAATGAGGCGAGAGAAATTTTACATGCGGCAAAACGGGCCCATC 1200
DB 1522 TTCCAGATCAATCTGGGAAATGAGGCGAGAGAAATTTTACATGCGGCAAAACGGGCCCATC 1581
QY 1201 AGTCCACCCCTGTGATGACACGCCCCATCAAAAGGGGGGAAATTCAGCTGAGACTG 1260
DB 1582 AGTCCACCCCTGTGATGACACGCCCCATCAAAAGGGGGGAAATTCAGCTGAGACTG 1641
QY 1261 GAAATGATCACTGTCAACACTGTCTCACTTCAAGAGGAGCTGCTGTATCCGATGCGG 1320
DB 1642 GAAATGATCACTGTCAACACTGTCTCACTTCAAGAGGAGCTGCTGTATCCGATGCGG 1701
QY 1321 ATATATGTGTGCGAGTACCATTC 1344
DB 1702 ATATATGTGTGCGAGTACCATTC 1725

RESULT 9
AA37670
ID AA37670 standard; cDNA: 2509 bp.

XX AAX37670;
AC 11-SEP-2000 (first entry)
DE Human EGF-like homologue (PRO217) encoded by DNA32279 cDNA.
XX
XX Inflammatory cell infiltration; immune response; T cell proliferation;
KW anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthritis;
KW T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
KW inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
KW diabetes mellitus; demyelinating polyneuropathy; hepatitis; cirrhosis; enteropathy;
KW multiple sclerosis; polynuropathy; inflammatory bowel disease; Whipple's disease;
KW skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
KW food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
KW idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
Egf-like; ss.
XX
XX Homo sapiens.
OS
PN W09914241-A2.
XX
XX 25-MAR-1999.
PD
PF 17-SEP-1998; 98WO-US19437.
XX
XX 17-SEP-1997; 97US-0059119.
PR 18-SEP-1997; 97US-0059263.
PR 28-OCT-1997; 97US-0063530.
PR 12-NOV-1997; 97US-0065186.
PR 21-NOV-1997; 97US-0066364.
PR 24-NOV-1997; 97US-0066770.
PR 04-JUN-1998; 98US-0088026.
XX
XX (GETH) GENENTECH INC.
XX
XX Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
PI WPI: 1999-229499/19.
XX
XX Composition containing novel polypeptide PRO245, its agonist or
PT antagonist -
XX
XX Example 1; Fig 5A-1-5A-2; 177pp; English.
XX
XX This invention describes a novel composition containing (part from a
CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
CC antagonist, or their fragments, for modulating: (i) infiltration of
CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
CC proliferation. The composition increases or decreases any of the effects
CC (i)-(iii). The products of the invention have anti-inflammatory,
CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
CC and their fragments, are used to treat immune-related diseases,
CC particularly T cell-mediated diseases. The diseases treated include
CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
CC arthritis, spondyloarthritis, systemic sclerosis (scleroderma),
CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
CC purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,
CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
CC thyroiditis), diabetes mellitus, immune-mediated renal disease
CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
CC inflammatory demyelinating polyneuropathy, infectious hepatitis
CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
CC (ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and
CC Whipple's disease. Autoimmune or immune-mediated skin diseases including
CC bullous skin diseases, erythema multiforme/contact dermatitis, psoriasis,

```

OY 361 GAAAGCAACCAATGTGTGATGTGGACGAGTGTGCAACAGATTCACCAAGTCCAAACCCC 420
DB 546 GAAAGCAACCAATGTGTGATGTGGACGAGTGTGCAACAGATTCACCAAGTCCAAACCCC 605
OY 421 ACCCAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCCGACACCGAGGATATTTGG 480
DB 606 ACCCAGATCTGCATCAATCTGAAGGCGGGTACACCTGCTCCGACACCGAGGATATTTGG 665
OY 481 CTTGCGGAAGGCGGATGCTTACATTTGATGAATGTGCGATGATGCTTACTCCAGCAGCTC 540
DB 666 CTTGCGGAAGGCGGATGCTTACATTTGATGAATGTGCGATGATGCTTACTCCAGCAGCTC 725
OY 541 TGTGCGAATGTCTCTGATCTCTATCTTCTTACATGCAACCCCTGGTTTACCTTCATATGAG 600
DB 726 TGTGCGAATGTCTCTGATCTCTATCTTCTTACATGCAACCCCTGGTTTACCTTCATATGAG 785
OY 601 GATGGAAGGCTTGTCCAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCACAAAC 660
DB 786 GATGGAAGGCTTGTCCAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGCACAAAC 845
OY 661 TGCGTCAACACCTACGCGCTCTTTCATCTGCGCGCTGTGACCCAGGATATGAATTTGAGGAA 720
DB 846 TGCGTCAACACCTACGCGCTCTTTCATCTGCGCGCTGTGACCCAGGATATGAATTTGAGGAA 905
OY 721 GATGCGCTTCAATGCAATATGACAGAGTGCAGCTTCTCTGATGCTTCTGCTCCACAT 780
DB 906 GATGCGCTTCAATGCAATATGACAGAGTGCAGCTTCTCTGATGCTTCTGCTCCACAT 965
OY 781 GAGTGTGTAAACAGCGCGGCAATCTTGTGCTCCGCGCTGTGACCGCTTACATCTGCTG 840
DB 966 GAGTGTGTAAACAGCGCGGCAATCTTGTGCTCCGCGCTGTGACCGCTTACATCTGCTG 1025
OY 841 GATGCAACCGAAGCTGTCCAAAGACATCAACGAATGTGAGACAGAGAACACACGTCGAC 900
DB 1026 GATGCAACCGAAGCTGTCCAAAGACATCAACGAATGTGAGACAGAGAACACACGTCGAC 1085
OY 901 CTGACAGCAGACGCTGTCAATTTTCAAGAGGGGCTTCAATTCATCGACCCCATCCGCTG 960
DB 1086 CTGACAGCAGACGCTGTCAATTTTCAAGAGGGGCTTCAATTCATCGACCCCATCCGCTG 1145
OY 961 GAGGAGCCTTATGTGAGGATCAGTATACCGCTGTATGTGCTGCTGAGAACCCCTGCG 1020
DB 1146 GAGGAGCCTTATGTGAGGATCAGTATACCGCTGTATGTGCTGCTGAGAACCCCTGCG 1205
OY 1021 TCGAGAGCAGACGCTTACATCTTGTACCGGAGATGACGCTGTGCTCAGGACGCTCC 1080
DB 1206 TCGAGAGCAGACGCTTACATCTTGTACCGGAGATGACGCTGTGCTCAGGACGCTCC 1265
OY 1081 GTTCCCGCTGTACATCTTCCAAATTCAGACGACGACCGCTACCCCTGGGGCCATTTACAT 1140
DB 1266 GTTCCCGCTGTACATCTTCCAAATTCAGACGACGACGACCGCTACCCCTGGGGCCATTTACAT 1325
OY 1141 TTCCAGATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCCCATC 1200
DB 1326 TTCCAGATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGGCAAAACGGGCCCCCATC 1385
OY 1201 AGTGCACACCTGTGTATGACACGCCCCATCAAAGGGCCCCGGGAAATTCAGCTGAGCTTG 1260
DB 1386 AGTGCACACCTGTGTATGACACGCCCCATCAAAGGGCCCCGGGAAATTCAGCTGAGCTTG 1445
OY 1261 GAATATGATCACTGTCAACAGCTGTCAATTCAGAGGAGCTCCGTGATCCGACCTGCGG 1320
DB 1446 GAATATGATCACTGTCAACAGCTGTCAATTCAGAGGAGCTCCGTGATCCGACCTGCGG 1505
OY 1321 ATATATGTGTGCGAGTACCCATTC 1344
DB 1506 ATATATGTGTGCGAGTACCCATTC 1529

```

RESULT 8
AAK05359
ID AAK05359 standard; DNA; 2550 BP.

```

XX AC AAK05359;
XX DT 07-MAY-1999 (first entry)
XX DE Human extracellular matrix protein (ECMP)-1 encoding DNA.
XX DE Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
XX DE Immune disorder; human; ss.
XX OS Homo sapiens.
XX PH Key
XX FT CDS
XX FT Location/Qualifiers
XX FT 382..1728
XX FT /tag="a"
XX FT /product="ECMP-1"
XX PN W09900410-A2.
XX PD 07-JAN-1999.
XX PE 23-JUN-1998; 98MO-US13012.
XX PR 27-JUN-1997; 97US-0884072.
XX PA (INCYTE PHARM INC.
XX PI Bandman O, Corley NC, Guegler KJ;
XX DR WPI; 1999-095674/08.
XX DR P-PSDB; AAW94281.
XX PT New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX PT useful in the diagnosis, prevention and treatment of immune
XX PT disorders and cancer
XX PS Claim 3; Fig 1A-G; 79pp; English.
XX CC This DNA encodes a human extracellular matrix protein (ECMP)-1. Host
XX CC cells containing a vector comprising the ECMP-1 nucleic acid are used
XX CC for the recombinant production of the protein. ECMP-1 and its
XX CC (ant)agonists, are useful in the diagnosis, prevention, and treatment
XX CC of cancer and immune disorders.
XX CS Sequence 2550 BP; 623 A; 690 C; 596 G; 641 T; 0 other;
SQ
Query Match 100.0%; Score 1344; DB 20; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 1344; Conservative 0; Indels 0; Gaps 0;
OY 1 ATCCAGGATTAAGGATTAATCACTGTTACATTTCTGCTCTGCTTCCAAACCCCT 60
DB 382 ATCCAGGATTAAGGATTAATCACTGTTACATTTCTGCTCTGCTTCCAAACCCCT 441
OY 61 GGAATGACAGCAGCAGTGCAGAAATGGCTTTGACCTGATCCGACGAGGAGAGT 120
DB 442 GGAATGACAGCAGCAGTGCAGAAATGGCTTTGACCTGATCCGACGAGGAGAGT 501
OY 121 TTTGATTTGATGTAATGCCAATCCATCCCGAGGCGCTCCGAGGAGACATGATGTGT 180
DB 502 TTTGATTTGATGTAATGCCAATCCATCCCGAGGCGCTCCGAGGAGACATGATGTGT 561
OY 181 AACCAAAATGGCGGATTTATGATTCACCGGACAAACCTGTGTATGAGGCGCTTC 240
DB 562 AACCAAAATGGCGGATTTATGATTCACCGGACAAACCTGTGTATGAGGCGCTTC 621
OY 241 TCGAACCCCTACTGACCCCTACTGACAGTCCCTACCGAGCGCGCCACACACTCTCA 300
DB 622 TCGAACCCCTACTGACCCCTACTGACAGTCCCTACCGAGCGCGCCACACACTCTCA 661
OY 301 GCTCAACATCTCCACAGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGAT 360
DB 682 GCTCAACATCTCCACAGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGAT 741

```


DB 1009 GATGACACCGAAGCTGCCAAGACATCAACGAATGTGAGACACAGAACACACGTCGAC 1068
 QY 901 CTGACGACACAGCTGCTACAAATTACAAAGGGGGCTTCAAAATGATCGACCCCTCGCTGT 960
 DB 1069 CTGACGACACAGCTGCTACAAATTACAAAGGGGGCTTCAAAATGATCGACCCCTCGCTGT 1128
 QY 961 GAGGAGCCCTTATCTGAGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGCG 1020
 DB 1129 GAGGAGCCCTTATCTGAGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGCG 1188
 QY 1021 TGCAGACACAGCCCTTATCTGAGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGCG 1080
 DB 1189 TGCAGACACAGCCCTTATCTGAGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGCG 1248
 QY 1081 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTCGGGGCTATTACAT 1140
 DB 1249 GTTCCCGCTGACATCTTCCAAATGCAAGCAGCAGCCGCTACCTCGGGGCTATTACAT 1308
 QY 1141 TTCCAGATCAATCTGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCGCCATC 1200
 DB 1309 TTCCAGATCAATCTGGAATGAGGGGAGAGAAATTTTACATGCGGCAAAAGGGGCGCCATC 1368
 QY 1201 AGTGCACCCCTGTGTGATGACACGCCCATCAAAAGGGGCGGGAATTCACCTGACCTTG 1260
 DB 1369 AGTGCACCCCTGTGTGATGACACGCCCATCAAAAGGGGCGGGAATTCACCTGACCTTG 1428
 QY 1261 GAAATATATCACTGATCAACATGTCATCACTTTCAGAGGACGCTCCGCTGATCCGATGCGG 1320
 DB 1429 GAAATATATCACTGATCAACATGTCATCACTTTCAGAGGACGCTCCGCTGATCCGATGCGG 1488
 QY 1321 ATATATGTGTGCGAGTACCCATTC 1344
 DB 1489 ATATATGTGTGCGAGTACCCATTC 1512

RESULT 6
 AAX07567
 ID AAX07567 standard; cDNA; 2362 BP.

XX AC AAX07567;
 XX DT 21-JUN-1999 (first entry)
 XX DE Homo sapiens fetal kidney clone AK647 secreted protein gene.
 XX KW Secreted protein; fetal kidney; ds.
 XX DS Homo sapiens.
 XX FT Key Location/Qualifiers
 FT CDS 186..1532
 FT /tag- a
 FT /product- "secreted protein"
 XX MO9900405-AI.
 XX PD 07-JAN-1999.
 XX PF 29-JUN-1998; 98MO-US13530.
 XX PR 30-JUN-1997; 97US-0885610.
 XX PA (GEMV) GENETICS INST INC.
 XX PI Agostino MJ, Evans C, Jacobs K, Lavallic ER, McCoy JM,
 XX Merberg D, Racie LA, Treacy M;
 XX DR MPI: 1999-095671/08.
 XX DR P-PSDB: AAM95709.
 XX PT New polynucleotides encoding secreted human proteins - are derived
 PT from foetal kidney or adult retina cDNA libraries, used as, e.g.
 PT potential vaccines

XX PS Disclosure: Pages 51-52; 76pp; English.

XX The sequence is that encoding a secreted protein from a human fetal
 CC kidney clone AK647. Such a sequence is predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals, although no
 CC supporting data is given. Suggested activities include nutritional
 CC activity, cytokine and cell proliferation/differentiation activity,
 CC haematopoiesis (e.g. as vaccines) or suppressing activity,
 CC CC and thrombolytic activity, chemotactic/chemokinetic activity, hemostatic
 CC activity, cadherin/tumour invasion suppressor activity, and tumour
 CC inhibition activity. It is also stated to be useful for gene
 CC therapy.

SQ Sequence 2362 BP; 580 A; 650-C; 548 G; 584 T; 0 other;

Query Match 100.0%; Score 1344; DB 20; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCAGGATTAAGAGATACATCTGATCAATCTGCTCTCTTCCAAAGCCT 60
 DB 186 ATGCCAGGATTAAGAGATACATCTGATCAATCTGCTCTCTTCCAAAGCCT 245
 QY 61 GGGATATGACAGCAGCAGTGCAGCAATGCTTTCACCTGATCCGCACTGAGCAGTGT 120
 DB 246 GGGATATGACAGCAGCAGTGCAGCAATGCTTTCACCTGATCCGCACTGAGCAGTGT 305
 QY 121 TTGATATTTGATGAATCCCAACCATCCCGAGGCTTCCAGAGAGATGATGTGT 180
 DB 306 TTGATATTTGATGAATCCCAACCATCCCGAGGCTTCCAGAGAGATGATGTGT 365
 QY 181 AACCAAAATGGCGGATTTATGATTCCTCCGGAACAACTGTGTATGAGGCGCTTAC 240
 DB 366 AACCAAAATGGCGGATTTATGATTCCTCCGGAACAACTGTGTATGAGGCGCTTAC 425
 QY 241 TCGAACCCCTACTGACACCCCTACTCAGTCCGATACCCAGACAGTCCGCCCACTGCA 300
 DB 426 TCGAACCCCTACTGACACCCCTACTCAGTCCGATACCCAGACAGTCCGCCCACTGCA 485
 QY 301 GCTCAAAATATCCACAGATCTCCAGGCTCTTATATGCTGCTTGTGATACAGATGAT 360
 DB 486 GCTCAAAATATCCACAGATCTCCAGGCTCTTATATGCTGCTTGTGATACAGATGAT 545
 QY 361 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCACCAAGTGCACCCC 420
 DB 546 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGCAACAGATTCACCAAGTGCACCCC 605
 QY 421 ACCCAAGTTCGATCAATGATGAGAGGCGGATACCTGCTGACAGGAGATATGAG 480
 DB 606 ACCCAAGTTCGATCAATGATGAGAGGCGGATACCTGCTGACAGGAGATATGAG 665
 QY 481 CTCTGGAAGGACAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 666 CTCTGGAAGGACAGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 725
 QY 541 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 726 TGTGCAATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
 QY 601 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 786 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
 QY 661 TGCCTCAACACCTACGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 846 TGCCTCAACACCTACGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
 QY 721 GATGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

OY	961	GAGAGGCTTATCTGAGATCAGGATTAACCGCTGTATGTCTCGTGGAACCCGTGGC	1020
Db	1166	GAGAGGCTTATCTGAGATCAGGATTAACCGCTGTATGTCTCGTGGAACCCGTGGC	1225
OY	1021	TGCAGAGACCAAGCCCTTACCATTTTGTACCGGACATGACCTGTCTCAGACGCTCC	1080
Db	1226	TGCAGAGACCAAGCCCTTACCATTTTGTACCGGACATGACCTGTCTCAGACGCTCC	1285
OY	1081	GTTTCCCGCTGACATCTTTCAAATGCAAGCCACGACCCGCTACCTGTGGGGCCTATTACATT	1140
Db	1286	GTTTCCCGCTGACATCTTTCAAATGCAAGCCACGACCCGCTACCTGTGGGGCCTATTACATT	1345
OY	1141	TTTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTATCAGTGGGCAAAACGGGGCCCATC	1200
Db	1346	TTTCCAGATCAAAATGTGGGAATGAGGGGAGAGAAATTTATCAGTGGGCAAAACGGGGCCCATC	1405
OY	1201	AGTCCCAACCCGTGTGATGACACGGCCCATCAAAAGGGCCCGGGAATTCAGACTGGACTTG	1260
Db	1406	AGTCCCAACCCGTGTGATGACACGGCCCATCAAAAGGGCCCGGGAATTCAGACTGGACTTG	1465
OY	1261	GAAATGATCACTGTTCACACACTGTCTCACTCAATTCAGAGGCACTCCGTGATCGACTGGGG	1320
Db	1466	GAAATGATCACTGTTCACACACTGTCTCACTCAATTCAGAGGCACTCCGTGATCGACTGGGG	1525
OY	1321	ATTATGTGTGCGAGTACCCATTTC	1344
Db	1526	ATTATGTGTGCGAGTACCCATTTC	1549

RESULT 5
AAZ40027
ID AAZ40027 standard; DNA; 2328 BP.

DT 15-FEB-2000 (first entry)

Full length human A55 protein coding sequence.

KM A55 protein; human; smooth muscle proliferation; tissue generation;
KM vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;
KM vascular endothelial thickening; haematopoietic cell-regulator; cytokine
KM percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
KM actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
KM tumour metastasis inhibitor; ss.

OS Homo sapiens

PN W09955864-A1

04-NOV-1999

PF 28-APR-1999; 99WO-JP02284.

PR 28-APR-1998; 98JP-0119731.

PA (ONOH) ONO PHARM CO LTD.

PI Honjo T, Tashiro K, Nakamura T,

DR WPI; 2000-038647/03.

XX

[illegible]

PS Claim 4, 5; Page 76-80; 87pp; Japanese.

CC This sequence encodes the human A55 protein of the invention. The protein
CC can be used for the treatment of diseases due to abnormal proliferation
CC of smooth muscle. The polypeptides can be used according their inhibition
CC of the proliferation of vascular smooth muscle cells, particularly in
CC treating arteriosclerosis or re-narrowing by vascular endothelial

CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
CC tissue generation/repairation activity, actin/inhibin activity, taxis
CC and chemotaxis activity, blood coagulation/thrombotic activity,
CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
CC tumour inhibition, and as a nutrient.
CC Note: The open reading frame of this sequence, as well as the mature
CC protein coding sequence are specifically claimed in claim 4.
XX
SQ Sequence 2328 BP, 560 A; 646 C; 541 G; 561 T; 0 other;

Query Match	100.0%	Score 1344	DB 21	Length 2328
Best Local Similarity	100.0%	Pred. NO. 0		
Matches 1344	0	Mismatches	0	Gaps 0

QY	1	ATGCCAGAAATAAAAAGATACTCACTGTATACATTTCTGGCTCTGTGCTTCCAAAGCCCT	60
Db	169	ATGCCAGAAATAAAAAGATACTCACTGTATACATTTCTGGCTCTGTGCTTCCAAAGCCCT	228
QY	61	GGAAATGCACAGGCGCAGTGCACGAATGGCTTTGACCTGAGATCCGACAGTACAGACATGT	120
Db	229	GGAAATGCACAGGCGCAGTGCACGAATGGCTTTGACCTGAGATCCGACAGTACAGACATGT	288
QY	121	TTAGATATTGATGAATGGCGGAACCAATCCCGAGGCGTCCGAGAGACATGATGTGTGTT	180
Db	289	TTAGATATTGATGAATGGCGGAACCAATCCCGAGGCGTCCGAGAGACATGATGTGTGTT	348
QY	181	AACCAAAATGGCGGTATTATTATGATTCCTCCGGAAACAAACCTGTATATCGAGGGCCCTAC	240
Db	349	AACCAAAATGGCGGTATTATTATGATTCCTCCGGAAACAAACCTGTATATCGAGGGCCCTAC	408
QY	241	TCGAACCCCTACTGCAACCCCTACTACAGATCCGTACCCAGCAAGTCCGCCACCACTCTCA	300
Db	409	TCGAACCCCTACTGCAACCCCTACTACAGATCCGTACCCAGCAAGTCCGCCACCACTCTCA	468
QY	301	GCTCCAAACTATCCACGATCTCCAGGCGCTTTATATGCGCGTTTGATATCCAGATGAT	360
Db	469	GCTCCAAACTATCCACGATCTCCAGGCGCTTTATATGCGCGTTTGATATCCAGATGAT	528
QY	361	GAAGCAACCAATGTGTGATGTGGAGAGTGTGCAACAGATTCACACAGTGCAAACCC	420
Db	529	GAAGCAACCAATGTGTGATGTGGAGAGTGTGCAACAGATTCACACAGTGCAAACCC	588
QY	421	ACCGAGATCTGATTAATACTGTAAGAGCGGGTACACCTGCTCTCGACCGACGAGATATTGG	480
Db	589	ACCGAGATCTGATTAATACTGTAAGAGCGGGTACACCTGCTCTCGACCGACGAGATATTGG	648
QY	481	CTTCTGGAAGGCCAGTGTACACTTTGATGATGTGCGTATGTTACTGCCAGACGCTC	540
Db	649	CTTCTGGAAGGCCAGTGTACACTTTGATGATGTGCGTATGTTACTGCCAGACGCTC	708
QY	541	TGTGGAGATGTTCCGAGATGCTATCTGTATGACATGCAACCCGGTTTAAACCTCAATGAG	600
Db	709	TGTGGAGATGTTCCGAGATGCTATCTGTATGACATGCAACCCGGTTTAAACCTCAATGAG	768
QY	601	GATGGAAGGTCTTGGCAAGATGTGAACAGATGTGCCACCGAAGAACCCCTGCGTGCAAAC	660
Db	769	GATGGAAGGTCTTGGCAAGATGTGAACAGATGTGCCACCGAAGAACCCCTGCGTGCAAAC	828
QY	661	TGCGTCAACACCTACGGCTCTTTTCACTTTCGCGCTGTGACCCAGATATGAACTTAGGAA	720
Db	829	TGCGTCAACACCTACGGCTCTTTTCACTTTCGCGCTGTGACCCAGATATGAACTTAGGAA	888
QY	721	GATGGCGTTCAATGCAAGTATATGAGAGAGTGCACCTCTGAGATTCCTCTGCCAACAT	780
Db	889	GATGGCGTTCAATGCAAGTATATGAGAGAGTGCACCTCTGAGATTCCTCTGCCAACAT	948
QY	781	GAGTGTGTGAACCAAGCCCGGACCAATACTTTCGCTTCGCGCTCCAGGGCTACATCCTGCTG	840
Db	949	GAGTGTGTGAACCAAGCCCGGACCAATACTTTCGCTTCGCGCTCCAGGGCTACATCCTGCTG	1008
QY	841	GATGCAACCGGAAGCTGCCAAGACATCAACGAATGTGACACAGGAAACCAACAGTGCAC	900

QY 1021 TGCAGACGACGCCCTTACCATCTGTACCGGGAGATGAGCGTGGTCAAGACGTC 1080
 DB 1231 TGCAGACGACGCCCTTACCATCTGTACCGGGAGATGAGCGTGGTCAAGACGTC 1290
 QY 1081 GTTCCCGTGCATCTTCCAAATGCAAGCCGCGCTACCTGGGGCTATTACAT 1140
 DB 1291 GTTCCCGTGCATCTTCCAAATGCAAGCCGCGCTACCTGGGGCTATTACAT 1350
 QY 1141 TTCCGACATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGCAACGGCCCATTC 1200
 DB 1351 TTCCGACATCAATCTGGGAATGAGGCGAGAAATTTTACATGCGCAACGGCCCATTC 1410
 QY 1201 ATGTCCACCTGGTGTGATGACAGCGCCCATCAAAAGGGCCCGGAAATCAGCTGGACTTG 1260
 DB 1411 ATGTCCACCTGGTGTGATGACAGCGCCCATCAAAAGGGCCCGGAAATCAGCTGGACTTG 1470
 QY 1261 GAAATGATCACTGTCAACACTGTCACTCACTGAGAGGACGCTCGTGTGATCGACTGGCG 1320
 DB 1471 GAAATGATCACTGTCAACACTGTCACTCACTGAGAGGACGCTCGTGTGATCGACTGGCG 1530
 QY 1321 ATATATGTGTGCGCATACCATTC 1344
 DB 1531 ATATATGTGTGCGCATACCATTC 1554
 RESULT 4
 AAK94505
 ID AAK94505 standard; cDNA; 2126 BP.
 AC AAK94505;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human full-length cDNA, SEQ ID NO: 3356.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 OS Homo sapiens.
 XX
 PN EP130094-A2.
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 (HELI-) HELIX RES INST.
 PA
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR MPI: 2001-524255/58.
 DR P-PSDB; AAM93573.
 PT
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3356; 1380bp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a full length
 CC human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX

SQ Sequence 2126 BP; 501 A; 608 C; 520 G; 497 T; 0 other;

Query Match
 Best Local Similarity 100.0%; Score 1344; DB 22; Length 2126;
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCAGGATTAATAAGATACACGCTTACCATCTGTACCGGGAGATGAGCGTGGTCAAGACGTC 60
 DB 206 ATGCCAGGATTAATAAGATACACGCTTACCATCTGTACCGGGAGATGAGCGTGGTCAAGACGTC 265
 QY 61 GGGAAATGACAGGACAGTGCAGCAAGATGCTTACCTGATGATGCGCAATGAGACAGTGT 120
 DB 266 GGGAAATGACAGGACAGTGCAGCAAGATGCTTACCTGATGATGCGCAATGAGACAGTGT 325
 QY 121 TTGATATATGATGAATGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGT 180
 DB 326 TTGATATATGATGAATGCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGT 385
 QY 181 AACCAAAATGGCGGTATTTATGATTCCTCCGGAACAAACCTGTGATGAGGGCCCTAC 240
 DB 386 AACCAAAATGGCGGTATTTATGATTCCTCCGGAACAAACCTGTGATGAGGGCCCTAC 445
 QY 241 TCGAACCCCTACTGACCCCTTACTAGTCCGATACCGAGAGCTGCGCCACACTCTCA 300
 DB 446 TCGAACCCCTACTGACCCCTTACTAGTCCGATACCGAGAGCTGCGCCACACTCTCA 505
 QY 301 GCTCCAAATATCCCAACGATCTCAAGGCTCTTATATGCCGCTTTGATACCATGAT 360
 DB 506 GCTCCAAATATCCCAACGATCTCAAGGCTCTTATATGCCGCTTTGATACCATGAT 565
 QY 361 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTCAACCC 420
 DB 566 GAAAGCAACCAATGTGTGATGTGAGAGAGTGTGCAAGATTTCCACAGTCAACCC 625
 QY 421 ACCCAGATCTGATCAATGATGAGAGGCGGTACACCTCTCTGACAGCAGCATATTTG 480
 DB 626 ACCCAGATCTGATCAATGATGAGAGGCGGTACACCTCTCTGACAGCAGCATATTTG 685
 QY 481 CTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 686 CTCTGGAAGGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 745
 QY 541 TGTGCGAATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 746 TGTGCGAATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 805
 QY 601 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 806 GATGGAAGTCTTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
 QY 661 TGCCTCAACACCTACGCGCTCTTCACTGCGCTGTGACCCAGATATCAACTGTGAGAA 720
 DB 866 TGCCTCAACACCTACGCGCTCTTCACTGCGCTGTGACCCAGATATCAACTGTGAGAA 925
 QY 721 GATGCGCTTCACTGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 926 GATGCGCTTCACTGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
 QY 781 GATGCTGTGACAGCCCGGACATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 986 GATGCTGTGACAGCCCGGACATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 QY 841 GATGACAAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 1046 GATGACAAACGAGCTGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1105
 QY 901 CTGAGACAGAGCTGTCAATTTCAAGGGGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 1106 CTGAGACAGAGCTGTCAATTTCAAGGGGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1165

ID ABR13627 standard; CDNA: 1720 BP.
 AC ABR13627;
 DT 23-APR-2002 (first entry)
 DE Human CDNA encoding extracellular protein-like/EGF-like protein, EGF.
 XX
 XX Human; ss; gene; extracellular protein-like protein; EGF-like;
 KM protein; epidermal growth factor; EGF; ATCC 97285; gene therapy;
 KM vascular smooth muscle cell proliferation; Marfan syndrome;
 KM wound healing; neurological trauma; acquired immunodeficiency syndrome;
 KM AIDS-related dementia; ocular disorder; kidney disorder; liver disorder;
 KW hair follicle growth promotion; burn; ulcer; corneal incision;
 KW corneal inflammation; neoplasm; psoriasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 211..1557
 FT /*tag= a
 FT /product= "EGF"
 FT /note= "This CDS minus the stop codon (211-2554) is
 FT specifically claimed in claim 5"
 FT 211..285
 FT /*tag= b
 FT mat_peptide 286..1554
 FT /*tag= c
 FT /label= Mature_EGF
 FT /note= "This region is specifically claimed in
 FT claim 4"
 XX
 PN US2001051358-A1.
 XX
 PD 13-DEC-2001.
 XX
 XX 25-MAR-1999; 990US-0275805.
 XX
 PR 11-APR-1997; 970US-0839525.
 PR 10-APR-1996; 96WO-US05247.
 XX
 PA (OLSE/) OLSEN H S.
 PA (LIHH/) LI H.
 PA
 PI Olsen HS, Li H;
 PI
 DR WPI; 2002-121417/16.
 DR P-PSDB; AAU75494.
 XX
 XX New nucleic acid encoding human extracellular/epidermal growth factor,
 XX useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
 XX also related polypeptides -
 PT
 PS Claim 4; Fig 1; 22pp; English.
 XX
 XX The invention relates to a novel polynucleotide which is at
 CC least 95% identical with a sequence (ATCC 97283) encoding mature human
 CC extracellular protein-like/epidermal growth factor (EGF)-like protein,
 CC EGF. Also included are the EGF EGF domains, a vector containing
 CC the polynucleotide, a host cell containing the vector, anti-EGF
 CC antibodies and antagonists of EGF. The polynucleotide is used for
 CC recombinant production of EGF. In gene therapy, as hybridisation probes,
 CC as antisense antagonists and for chromosome identification. The protein
 CC is used to treat patients who require EGF, to identify specific
 CC antagonists, used to treat conditions that require inhibition of EGF
 CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
 CC healing, neurological trauma, acquired immunodeficiency syndrome
 CC (AIDS)-related dementia, ocular disorders, kidney disorders, liver
 CC disorders, hair follicle growth promotion, burns, ulcers, corneal
 CC incisions, corneal inflammation, neoplasm and psoriasis), to raise
 CC specific antibodies and to characterise receptors. The present
 CC sequence encodes EGF.
 XX

Seq	Sequence	1720 Bp	396 A	515 C	434 G	372 T	3 other:
50	Query Match	100.0%	Score 1344	DB 24	Length 1720		
	Best Local Similarity	100.0%	Prod. No. 0				
	Matches 1344	Conservative 0	Mismatches 0	Indels 0	Gaps 0		
QY	1	ATGCCAGGAATAAAAGATATCTCAGTGTACCAATTCCTGGGCTCTCTCTCTCCAAAGCCCT	60				
DB	211	ATGCCAGGAATAAAAGATATCTCAGTGTACCAATTCCTGGGCTCTCTCTCTCCAAAGCCCT	270				
QY	61	GGGATGCGACAGGACAGTGCACGAATGGCTTTTACCTGCATGCCACAGTCAGACAGTGT	120				
DB	271	GGGATGCGACAGGACAGTGCACGAATGGCTTTTACCTGCATGCCACAGTCAGACAGTGT	330				
QY	121	TTAGATATTTGATGTAATGCCCAACCATCCCGAGGCTTCGCGAGAGACATATGTGTGT	180				
DB	331	TTAGATATTTGATGTAATGCCCAACCATCCCGAGGCTTCGCGAGAGACATATGTGTGT	390				
QY	181	AACCAAAATGCGGGGTATTTTATGCAATTCGCCGGAACAAACCTGTGTATCGAGGCGCTTAC	240				
DB	391	AACCAAAATGCGGGGTATTTTATGCAATTCGCCGGAACAAACCTGTGTATCGAGGCGCTTAC	450				
QY	241	TCGAACCCCTACTGACGCCCTTACTCAGGTCCGTACCCAGCAGTGTGCCCACTCTCA	300				
DB	451	TCGAACCCCTTCTGACGCCCTTACTCAGGTCCGTACCCAGCAGTGTGCCCACTCTCA	510				
QY	301	GCTCCAAACATATCCAGCATCTCCAGGCGCTTTATATGCGCTTTGGTATACAGTGGAT	360				
DB	511	GCTCCAAACATATCCAGCATCTCCAGGCGCTTTATATGCGCTTTGGTATACAGTGGAT	570				
QY	361	GAAGCAACCAATGTGTGATGTGGACAGTGTGCAACAGATTTCCACCACTGCAACCCC	420				
DB	571	GAAGCAACCAATGTGTGATGTGGACAGTGTGCAACAGATTTCCACCACTGCAACCCC	630				
QY	421	ACCAGATCTGCATCAATACTGGAAGGGGGGTACACCTGCTCTGACCGAGGATATTGG	480				
DB	631	ACCAGATCTGCATCAATACTGGAAGGGGGGTACACCTGCTCTGACCGAGGATATTGG	690				
QY	481	CTTCTGGAAGCGCACAGTGTGTAGACATGTGATCAATGTGCTATGTGTACTGCGACAGCTC	540				
DB	691	CTTCTGGAAGCGCACAGTGTGTAGACATGTGATCAATGTGCTATGTGTACTGCGACAGCTC	750				
QY	541	TGTGCGAATGTCTTCGGAATCCTATTTCTTTATATGCAACCCCTGGTTTACCCTCAATGAG	600				
DB	751	TGTGCGAATGTCTTCGGAATCCTATTTCTTTATATGCAACCCCTGGTTTACCCTCAATGAG	810				
QY	601	GATGAAGAGTCTTGCCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGGTGTCAAAAC	660				
DB	811	GATGAAGAGTCTTGCCCAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGGTGTCAAAAC	870				
QY	661	TGCGTCAACACCTTACGCGCTCTTTCATCTGCGGTGTGACCCAGGATATGAACCTTGAGAA	720				
DB	871	TGCGTCAACACCTTACGCGCTCTTTCATCTGCGGTGTGACCCAGGATATGAACCTTGAGAA	930				
QY	721	GATGGCGCTTCAATGCAAGTATATGAGAGAGTGCACAGCTCTCTGTAGTTCTCTGCGCAACAT	780				
DB	931	GATGGCGCTTCAATGCAAGTATATGAGAGAGTGCACAGCTCTCTGTAGTTCTCTGCGCAACAT	990				
QY	781	GAGTGTGTGAACCAAGCCCGGACACATATTCTGCTCTGCGCCCTCTCAGGCTACATCTGTG	840				
DB	991	GAGTGTGTGAACCAAGCCCGGACACATATTCTGCTCTGCGCCCTCTCAGGCTACATCTGTG	1050				
QY	841	GATGACAAACCAAGCTGTGCACAGACATCAACGAATGTGAGACACAGAGAACCAAGCTGCAAC	900				
DB	1051	GATGACAAACCAAGCTGTGCACAGACATCAACGAATGTGAGACACAGAGAACCAAGCTGCAAC	1110				
QY	901	CTGAGACGAGAGTGTCTACAAATTTACAAGGGGGGCTTCAAAATGCATGCAACCCCATCGGTGT	960				
DB	1111	CTGAGACGAGAGTGTCTACAAATTTACAAGGGGGGCTTCAAAATGCATGCAACCCCATCGGTGT	1170				
QY	961	GAGGAGCCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCTGTGGC	1020				
DB	1171	GAGGAGCCTTATCTGAGGATCAGTATTAACCGCTGTATGTCTCTGCTGAGAACCTGTGGC	1230				

note: Extracellular/epidermal growth factor-like"

XX MO9846746-1
XX 22-OCT-1998.
XX
XX 11-APR-1997; 97WO-US06020.
XX
XX 11-APR-1997; 97WO-US06020.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX L1 H, Olsen HS:
XX WPI: 1998-568728/48.
XX P-PSDB: AAW79739.
XX
XX New Isolated extracellular/epidermal growth factor - used for
XX regulating vascular smooth muscle cell proliferation, e.g. for
XX enhancing neurological functions or treating neoplasia and other
XX disorders.

XX Claim 1a; Fig 1A-D; 62pp; English.

XX This sequence encodes a novel human extracellular/epidermal growth
XX factor-like protein, EGF. This protein can be used to regulate
XX vascular smooth muscle cell proliferation and for restoration or
XX enhancement of neurological functions diminished as a result of other
XX damaging pathologies such as AIDS dementia. The protein can also be used
XX to treat senile dementia, ocular disorders such as corneal inflammation,
XX for targeting tumor cells, for treating kidney disorders, for liver
XX regeneration or treating liver dysfunction, for treating wounds including
XX all cutaneous wounds, corneal wounds, and injuries to the
XX epithelial-lined hollow organs of the body or resulting from trauma such
XX as burns, abrasions and cuts as well as from surgical procedures such as
XX surgical incisions and skin grafting. The polypeptides can also be used
XX for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
XX other non-healing (trophic) conditions, to treat Marfan syndrome, to
XX promote hair follicular development, to stimulate growth and in
XX vitro and in vivo differentiation of various epidermal and epithelial cells in vitro and in
XX vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
XX treat neoplasia such as cancers or tumours, skin disorders such as
XX psoriasis or corneal inflammation. The products can also be used for
XX identifying EGF receptors, detection, diagnosis and drug screening.

XX Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other:

XX Query Match 100.0%; Score 1344; DB 19; Length 1720;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCCAGATATATAAGATGACTGTTACCTGTTGGCTCTGCTCTTCCAAAGCCCT 60
DB 211 ATGCCAGATATATAAGATGACTGTTACCTGTTGGCTCTGCTCTTCCAAAGCCCT 270
OY 61 GGGATGACAGGACAGTGCAGAAATGGCTTGACCTGATGCCCAATCAGACAGTGT 120
DB 271 GGGATGACAGGACAGTGCAGAAATGGCTTGACCTGATGCCCAATCAGACAGTGT 330
OY 121 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 331 TTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
OY 181 AACCAAAATGGCGGGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 391 AACCAAAATGGCGGGATTTTATGATGATGATGATGATGATGATGATGATGATGAT 450
OY 241 TCGAACCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGACAG 300
DB 451 TCGAACCCCTACTGACAGCCCTACTGACAGCCCTACTGACAGCCCTACTGACAG 510
OY 301 GGTCCAAACTATCCAGATCTCCAGGCTCTTATGATGATGATGATGATGATGATGAT 360
DB 511 GGTCCAAACTATCCAGATCTCCAGGCTCTTATGATGATGATGATGATGATGATGAT

DB 511 GGTCCAAACTATCCAGATCTCCAGGCTCTTATGATGATGATGATGATGATGATGAT 570
OY 361 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGACAGAGTGTGACAGAGTGTGACAG 420
DB 571 GAAAGCAACCAATGTGTGATGTGACAGAGTGTGACAGAGTGTGACAGAGTGTGACAG 630
OY 421 ACCCAGATCTGATCAATACGTAAGAGGCGGTACCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 631 ACCCAGATCTGATCAATACGTAAGAGGCGGTACCTGCTGCTGCTGCTGCTGCTGCTG 690
OY 481 CTTGTGAAGGCGCACTGCTTACATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 691 CTTGTGAAGGCGCACTGCTTACATGATGATGATGATGATGATGATGATGATGATGATG 750
OY 541 TGTGGAATGTTCTGCGATGCTTATGCTTGTATGATGATGATGATGATGATGATGATG 600
DB 751 TGTGGAATGTTCTGCGATGCTTATGCTTGTATGATGATGATGATGATGATGATGATG 810
OY 601 GATGGAAGGCTTGGCAAGATGTAAGAGAGTGTGACAGAGTGTGACAGAGTGTGACAG 660
DB 811 GATGGAAGGCTTGGCAAGATGTAAGAGAGTGTGACAGAGTGTGACAGAGTGTGACAG 870
OY 661 TGGTCAACACCTAGCGCTCTTTCATCTGCGCGCTGTGACAGAGTGTGACAGAGTGTG 720
DB 871 TGGTCAACACCTAGCGCTCTTTCATCTGCGCGCTGTGACAGAGTGTGACAGAGTGTG 930
OY 721 GATGCGCTTCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 931 GATGCGCTTCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 990
OY 781 GATGCTGTGTAACAGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 991 GATGCTGTGTAACAGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
OY 841 GATGACACCGGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 1051 GATGACACCGGAAGCTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1110
OY 901 GTTCAGACAGCGGTGCTTACAAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 960
DB 1111 GTTCAGACAGCGGTGCTTACAAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 1170
OY 961 GAGGAGCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
DB 1171 GAGGAGCCTTATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1230
OY 1021 TGCAGAGACAGCGCTTACCAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 1080
DB 1231 TGCAGAGACAGCGCTTACCAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 1290
OY 1081 GTTCAGACAGCGGTGCTTACCAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 1140
DB 1291 GTTCAGACAGCGGTGCTTACCAATGCAAGGCGGCTTCAAAATGCAATGCAATGCAATG 1350
OY 1141 TTCCAGATCAATCTGGAATGAGAGGAGAGATTTTATGATGATGATGATGATGATGAT 1200
DB 1351 TTCCAGATCAATCTGGAATGAGAGGAGAGATTTTATGATGATGATGATGATGATGAT 1410
OY 1201 AGTGCACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1411 AGTGCACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
OY 1261 GAAATGATCACTGTCAACACTGTGATCAATCTTCAAGAGGAGAGTCCGTGATCCAGTGG 1320
DB 1471 GAAATGATCACTGTCAACACTGTGATCAATCTTCAAGAGGAGAGTCCGTGATCCAGTGG 1530
OY 1321 ATATATGTGTGCAAGTACCATTC 1344
DB 1531 ATATATGTGTGCAAGTACCATTC 1554

RESULT 3
ABK13627

PT muscle cells expressing cell proliferating growth factor and celluia
 XX adherence factor -
 PS Example 2; Page 97; 98pp; English.

CC The present invention describes an artificial vascular graft (I
CC comprising a synthetic tubular element having a luminal surface coated
CC with several endothelial cells (ECs) and/or smooth muscle cells (SMCs)
CC genetically transformed to express at least one cell proliferating growth
CC factor (CGF), and a cellular adherence factor (CAF). (1) is useful for
CC replacing or bypassing a portion of a vascular system of an individual,
CC by implanting (1) into the vascular system of the individual, to form a
CC fluid communication between the vascular system and the graft, where the
CC synthetic tubular element of the graft is of an inner cross sectional
CC area which is substantially equivalent to an inner cross sectional
CC area of a blood vessel. The present sequence represents the human UP50 cDNA
CC sequence (also known as a developmental arteries and neural crest
CC epidermal growth factor (EGF)-like protein (DANCE) gene), which is used
CC in an example from the present invention.

Sequence 1347 BP; 329 A; 388 C; 327 G; 303 T; 0 other;

Query Match	100.0%;	Score 1344;	DB 24;	Length 1347;
-------------	---------	-------------	--------	--------------

Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	ATGCAGAAATAAAAGATCTACACCTTTCACATCTTGCGCTCTGTCTCCAGCCCT	60
OY	1	ATGCAGAAATAAAAGATCTACACCTTTCACATCTTGCGCTCTGTCTCCAGCCCT	60
Db	1	ATGCAGAAATAAAAGATCTACACCTTTCACATCTTGCGCTCTGTCTCCAGCCCT	60
OY	61	GGGAATGCACAGGCACAGTGCAGAAATGGCTTTCAGCTGGATTCGCACTCAGCAGCT	120
Db	61	GGGAATGCACAGGCACAGTGCAGAAATGGCTTTCAGCTGGATTCGCACTCAGCAGCT	120
OY	121	TTAGATATGATGAAATGCGCAACCATCCCGAGGCTCCGACGAGACATATGTGTGT	180
Db	121	TTAGATATGATGAAATGCGCAACCATCCCGAGGCTCCGAGGAGACATGATGTGTGT	180
OY	181	AACCAAAATGGCGGGTATTTATGCAATTCGCCGAGCAAAACCTGTGTATCGAGGGCCCTAC	240
Db	181	AACCAAAATGGCGGGTATTTATGCAATTCGCCGAGCAAAACCTGTGTATCGAGGGCCCTAC	240
OY	241	TCGAACCCCTACTGACACCCCTACTCAGATCCGTCACCGACAGCTGCCCCACACTCTCA	300
Db	241	TCGAACCCCTACTGACACCCCTACTCAGATCCGTCACCGACAGCTGCCCCACACTCTCA	300
OY	301	GCTCCAAACTATCCACAGATCTCCAGGCGCTTATATACCGCTTGTGATACCAATGGAT	360
Db	301	GCTCCAAACTATCCACAGATCTCCAGGCGCTTATATACCGCTTGTGATACCAATGGAT	360
OY	361	GAAAGCAACCAATGTGTGATGTGGAGAGATGTGCAACAGATTCCGACCACTGCAACCCC	420
Db	361	GAAAGCAACCAATGTGTGATGTGGAGAGATGTGCAACAGATTCCGACCACTGCAACCCC	420
OY	421	ACCAGATCTGCATCAATACTGAAGGGGGGTACACTGCTCCTGACCGAGGAGATATGG	480
Db	421	ACCAGATCTGCATCAATACTGAAGGGGGGTACACTGCTCCTGACCGAGGAGATATGG	480
OY	481	CTTCTGAGAGCCACAGTGGCTTAAGACATTTGATGATGTGCTATGTACTGCCAGAGCTC	540
Db	481	CTTCTGAGAGCCACAGTGGCTTAAGACATTTGATGATGTGCTATGTACTGCCAGAGCTC	540
OY	541	TGTGGAAATGTTCCGAGATCCTATCTTGTATCATCAACCCGTGTTTAACTTCATATGAG	600
Db	541	TGTGGAAATGTTCCGAGATCCTATCTTGTATCATCAACCCGTGTTTAACTTCATATGAG	600
OY	601	GATGAAAGCTTTCGCAAGATGTGAACAGATGTGACACGAGAAACCCCTGCGTGCACAC	660
Db	601	GATGAAAGCTTTCGCAAGATGTGAACAGATGTGACACGAGAAACCCCTGCGTGCACAC	660
OY	661	TGCGCAACACCTACGCGCTTTTCACTTCGCGCTGTGACCCAGATATGAATCTTAGGA	720
Db	661	TGCGCAACACCTACGCGCTTTTCACTTCGCGCTGTGACCCAGATATGAATCTTAGGA	720

OY		721	GATGGCCGTTCAATTGCAGATGATATGACAGATGACAGCTTCCTGTAGTTCCTCCTGCCAACAAT	780
Dd		721	GATGGCCGTTCAATTGCAGATGATATGACAGATGACAGCTTCCTGTAGTTCCTCCTGCCAACAAT	780
OY		781	GAGTGTTGAACAACGAGCCCCGGCACATACTTCTGCTCTGCTCCAGGCTACATCCTGCTG	840
Dd		781	GAGTGTTGAACAACGAGCCCCGGCACATACTTCTGCTCTGCTCCAGGCTACATCCTGCTG	840
OY		841	GATGATCAACCGAACACTGCGCAAGACATCATGCAATGTGAGCAGCAGAAACACACACTGCGAAC	900
Dd		841	GATGATCAACCGAACACTGCGCAAGACATCATGCAATGTGAGCAGCAGAAACACACACTGCGAAC	900
OY		901	CTGCAGCAGACGTCCTCAATTTTACAAAGGGGGCTTCAAATGTCATCGAACCCCATCGCTGT	960
Dd		901	CTGCAGCAGACGTCCTCAATTTTACAAAGGGGGCTTCAAATGTCATCGAACCCCATCGCTGT	960
OY		961	GAGGAGCCCTTAATTCGAGAGTACGTGATTAACCGCTGTATGTGTCCTGCTGAGAACCTTGCC	1020
Dd		961	GAGGAGCCCTTAATTCGAGAGTACGTGATTAACCGCTGTATGTGTCCTGCTGAGAACCTTGCC	1020
OY		1021	TGCGAGACACCGCCCTTTACATCTTGTGTACCGGGACATGACCGTGTGTACGAGACGCTCC	1080
Dd		1021	TGCGAGACACCGCCCTTTACATCTTGTGTACCGGGACATGACCGTGTGTACGAGACGCTCC	1080
OY		1081	GTTCCCGCGTGACATCTTTCACAAATGCACAGCCGCGCTACCTCGGGGCGCTATTACATT	1140
Dd		1081	GTTCCCGCGTGACATCTTTCACAAATGCACAGCCGCGCTACCTCGGGGCGCTATTACATT	1140
OY		1141	TTCCAGATCAAAATCTGGGAAATGAGGGGACAGAAATTTTACATGCGGCAAAACGGGGCCCATC	1200
Dd		1141	TTCCAGATCAAAATCTGGGAAATGAGGGGACAGAAATTTTACATGCGGCAAAACGGGGCCCATC	1200
OY		1201	AGTCCACACCTGGTGATGTACACCGCCCATCAAAAGGCCCCGGGAAATCCACACTGGACTTG	1260
Dd		1201	AGTCCACACCTGGTGATGTACACCGCCCATCAAAAGGCCCCGGGAAATCCACACTGGACTTG	1260
OY		1261	GAATATGATCACTGTCACACACTGTGTCATCAACTTCAGAGGACAGCTCCGTATCCGACTGCGG	1320
Dd		1261	GAATATGATCACTGTCACACACTGTGTCATCAACTTCAGAGGACAGCTCCGTATCCGACTGCGG	1320
OY		1321	ATATATGTGTGCGCAGTACCCATTTC	1344
Dd		1321	ATATATGTGTGCGCAGTACCCATTTC	1344
RESULT 2				
AAV62432				
ID	AAV62432	standard; DNA: 1720 BP.		
XX	AAV62432;			
AC				
XX	25-JAN-1999	(first entry)		
DT				
XX				
DE	Human EEGF genomic DNA.			
XX				
KW	Extracellular/epidermal growth factor-like protein; EGF; human; liver;			
KW	vascular smooth muscle cell proliferation; neurology; pathology; AIDS;			
KW	dementia; ocular; disorder; cornea; inflammation; tumour cell; kidney;			
KW	wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;			
KW	Marian syndrome; treatment; hair follicle; embryogenesis; neoplasma;			
KW	epidermal cell; cancer; psoriasis; detection; ss.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	211..1557		
FT		/*tag= a		
FT	sig_peptide	211..285		
FT		/*tag= b		
FT	mat_peptide	286..1557		
FT		/*tag= c		
FT	/product= "EEGF"			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 299.483 Seconds
(without alignments)
10106.379 Million cell updates/sec

Sequence: 1 atgcccagagataaaagat.....atgttcgcagctaccatc 1344

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 segs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database # 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1347	24	ABL52528
2	1344	100.0	1720	19	AAV62432
3	1344	100.0	1720	24	ABK13627
4	1344	100.0	2126	22	AAK94505
5	1344	100.0	2328	21	AAZ40027
6	1344	100.0	2362	20	AAZ07567
7	1344	100.0	2362	21	AAZ39892
8	1344	100.0	2550	20	AAZ05359
9	1342.4	99.9	2509	20	AAZ37670

10	1342.4	99.9	2609	22	AAZ46128
11	1342.4	99.9	2609	22	AAZ46128
12	1207.2	89.8	1228	18	AAZ89380
13	1113.6	82.9	1344	21	AAZ39383
14	1113.6	82.9	2233	21	AAZ39384
15	1113.6	82.9	2233	21	AAZ40029
16	1098.6	81.7	1383	21	AAZ39386
17	1098.6	81.7	2429	21	AAZ39387
18	1098.6	81.7	2429	21	AAZ40031
19	1056.2	78.6	1269	21	AAZ39385
20	1056.2	78.6	1269	21	AAZ39388
21	843.8	62.8	1254	21	AAZ61672
22	843.8	62.8	1254	22	AAZ99605
23	841.8	62.6	1260	24	ABL34757
24	776.4	57.8	1778	20	AAZ37673
25	721.4	53.7	853	22	AAZ31596
26	721.4	53.7	853	22	AAZ3159
27	721.4	53.7	1538	22	AAZ94129
28	636.4	47.4	680	22	AAZ44245
29	390	29.0	670	22	AAZ92195
30	390	29.0	670	22	AAZ93463
31	374.4	27.9	377	20	AAZ37686
32	371.2	27.6	1480	21	AAZ14608
33	371.2	27.6	1525	18	AAZ8974
34	371.2	27.6	1531	20	AAZ60351
35	371.2	27.6	1707	22	AAZ13771
36	371.2	27.6	1874	22	AAZ58583
37	371.2	27.6	1875	22	AAZ51402
38	371.2	27.6	1875	24	AAZ40256
39	371.2	27.6	1994	21	AAZ18229
40	371.2	27.6	2018	21	AAZ39800
41	347.2	25.8	1513	21	AAZ14603
42	337.6	25.1	1358	21	AAZ14616
43	326	24.3	354	20	AAZ37683
44	317	23.6	359	20	AAZ86404
45	316.6	23.6	1018	21	AAZ61786

ALIGNMENTS

RESULT 1
ABL52528
ID ABL52528 standard; cDNA; 1347 BP.
AC ABL52528;
DT 19-JUL-2002 (first entry)
DE Human UP50 cDNA sequence SEQ ID NO:1.
XX
XX Human; UP50; artificial vascular graft; cellular adherence factor;
KW cell proliferating growth factor; DANCE; epidermal growth factor; EGF;
KW developmental arteries and neural crest EGF-like protein; gene; ss.
XX
XX Homo sapiens.
OS
PN WO200207646-A2
XX
PD 31-JAN-2002.
XX
XX 20-JUL-2001; 2001MO-IL00670.
XX
XX 20-JUL-2000; 2000US-0620227.
XX
XX (MGVS-) MGVS LTD.
XX
XX Flugelman M, Preiss M, Gluzman Z, Koren B, Weisz A, Cohen T;
XX WPI; 2002-361632/39.
XX
XX Artificial vascular graft for therapy of vascular diseases, has
PT synthetic tubular element having a surface coated with endothelial,


```
Db 632 GATTTATGATTCCTCCGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTC 691
QY 423 GACCCCTACTCAGAGTCCGTACCCAGAGCTGCCACCACTCTCAGCTCCAACTATCC 482
Db 692 GACCCCTACTCAGAGTCCGTACCCAGAGCTGCCACCACTCTCAGCTCCAACTATCC 751
QY 483 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAAGATGATGAAGCAACCAATG 542
Db 752 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAAGATGATGAAGCAACCAATG 811
QY 543 TGTGATGTGAGAGTGTGCACACAGATCCCAAGTGCACACCCAGATCTGCAT 602
Db 812 TGTGATGTGAGAGTGTGC -AMAGATTCACACAGTGCAMCCACCCAGATCTGCAT 870
QY 603 CAATACGAAGGGGGGTACACCTGTCTGTGACCGAGGATATGGCTTGTGAAGGCCA 662
Db 871 CAATACGAAGGGGGGTACACCTGTCTGTGACCGAGGATATGGCTTGTGAAGGCCA 930
QY 663 GTGCTTAGACATGATGAATGTGCTATGGTTACTGCGACAGCTCTGTGGAATGTTC 722
Db 931 GTGCTTAGACATGATGAATGTGCTATGGTTACTGCGACAGCTCTGTGGAATGTTC 990
QY 723 TGGATCTATCTTGT 738
Db 991 TGGATCTATCTTGT 1006
```

Search completed: July 3, 2003, 17:36:06
Job time : 3436.81 secs

/clone="IMAGE:5229519"
 /clone.lib="NIH_MGC_120"
 /lab.host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 kb, insert size range 1-2.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."

BASE COUNT 184 a 240 c 195 g 175 t
 ORIGIN

Query Match 30.1%; Score 700.4; DB 13; Length 794;
 Best Local Similarity 98.4%; Pred. No. 4.8e-185;
 Matches 781; Conservative 0; Mismatches 6; Indels 7; Gaps 7;

28 CACGACTCGCTCGCCCTCTGGAATAAACACCCGAGCCCGAGGGCCAGAGAG 87
 1 CACGACTCGCTCGG -CCCTCTGGAATAAACACCCGAGCCCGAGGGCCAGAGAG 59
 88 CCGAGCTGCCGAGCTCTCGGGGGTCCGCCCGC -GAGCTTCTTCTCGCTTCGAT 146
 60 CCGAGCTGCCGAGCTCTCGGGGGTCCGCCCGC GAGCTTCTTCTCGCTTCGAT 119
 147 CTCTCTCTCGGGGCTTGGACATGCCAGGAATAAAGATGACTGTTACCATCT 206
 120 CTCTCTCTCGGGGCTTGGACATGCCAGGAATAAAGATGACTGTTACCATCT 179
 207 GGCCTCTCTCTTCCAGCCCTGGGAATGCACAGGCAAGTGCAGATGGCTTGCAC 266
 180 GGCCTCTCTCTTCCAGCCCTGGGAATGCACAGGCAAGTGCAGATGGCTTGCAC 239
 267 GGATGCCGATGAGGACAGTGTATGATGATGATGATGATGATGATGATGATGAT 326
 240 GGATGCCGATGAGGACAGTGTATGATGATGATGATGATGATGATGATGATGAT 299
 327 CCGAGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
 300 CCGAGGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
 387 CCCTCTGTATCGAGGGCCCTACTCGAACCCTACTCGAACCCTACTCGAAGTCCGTAC 446
 360 CCCTCTGTATCGAGGGCCCTACTCGAACCCTACTCGAACCCTACTCGAAGTCCGTAC 419
 447 AACAGTGGCCCACTCTCAGCTCCAAACTAT -CCGAGATCTCCAGGCTCTTATAT 505
 420 AACAGTGGCCCACTCTCAGCTCCAAACTAT -CCGAGATCTCCAGGCTCTTATAT 479
 506 GCGGCTTTGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
 480 GCGGCTTTGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 539
 566 CAGATTCACACAGTGC -AACCCACCCAGATCTCATCTCATCTCATCTCATCTCAT 624
 540 CAGATTCACACAGTGC -AACCCACCCAGATCTCATCTCATCTCATCTCATCTCAT 599
 625 TGCTCTCTGACGAGCAGATATTGGCTTCTGGAAGGCCAGTCTTAGACATTGATGAT 684
 600 TGCTCTCTGACGAGCAGATATTGGCTTCTGGAAGGCCAGTCTTAGACATTGATGAT 659
 685 CG -CTATGCTTACTGACAGAG -CTGTGTGGAATGTT -CCTGATCTCTATCTTGTACA 741
 660 CGCCTATGCTTACTGACAGAGCTGTGTGGAATGTT -CCTGATCTCTATCTTGTACA 719
 742 TGCAACCCCTGATTTTACCTCAATGAGAGTCTTGCCAGATGATGAGAGTGT 801
 720 TGCAACCCCTGATTTTAACTCAATGAGAGTCTTGCCAGATGATGAGAGTGT 779
 802 GCCACCGAGAACCC 815
 ||||||||| ||

Db 780 GCCACCGAGAGACC 793

RESULT 15
 AL545549
 LOCUS
 DEFINITION
 AL545549 LIT.NFL006.PL2 Homo sapiens cDNA clone CS001015YE20 5
 prime, mRNA sequence.
 ACCESSION
 AL545549
 VERSION
 AL545549.1 GI:12878031
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1 (bases 1 to 1006)
 Li, W.B., Gruber, C., Jeesee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
 source
 1..1006
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS001015YE20"
 /clone.lib="LIT.NFL006.PL2"
 /issue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
 was primed with a NotI-Oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact: Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com url :
 http://fulllength.invitrogen.com"

BASE COUNT 215 a 300 c 258 g 228 t 5 others
 ORIGIN

Query Match 30.1%; Score 700.4; DB 9; Length 1006;
 Best Local Similarity 98.2%; Pred. No. 5.2e-185;
 Matches 723; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

3 CCGGCGCTCTCCCGGTGTCTCTTCACGACTGCTGGCCCTCTGGAATAAACACC 62
 273 CCGGCGCTCTCCCGGTGTCTCTTCACGACTGCTGGCCCTCTGGAATAAACACC 332
 63 GCGAGCCCGGAGGCGCAGAGAGAGCGGAGGCGGAGGCTCTCGGGGGTCCGCCCG 122
 333 GCGAGCCCGGAGGCGCAGAGAGAGCGGAGGCGGAGGCTCTCGGGGGTCCGCCCG 392
 123 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGTGACATGCCAGATATA 182
 393 CGAGCTTCTTCTCGCTTGCATCTCTCTCGGCGCTTGTGACATGCCAGATATA 452
 183 AAGGATACGACTGTTACCATTTGCGCTCTGTCTTCCAGCCCTGGGATGACAGGC 242
 453 AAGGATACGACTGTTACCATTTGCGCTCTGTCTTCCAGCCCTGGGATGACAGGC 512
 243 AAGGATACGACTGTTACCATTTGCGCTCTGTCTTCCAGCCCTGGGATGACAGGC 302
 513 AAGGATACGACTGTTACCATTTGCGCTCTGTCTTCCAGCCCTGGGATGACAGGC 572
 303 ATGCCGAAACATCCCGAGGCTCCGAGAGAGATGATGATGATGATGATGATGATGAT 362
 573 ATGCCGAAACATCCCGAGGCTCCGAGAGAGATGATGATGATGATGATGATGATGAT 631
 363 GATTTATGATTCCTCCCGAGAACCTGTGTATGAGAGGCGCTTACTGACACCTCTACT 422
 ||||||||| ||

LOCUS BM974454 735 bp mRNA linear EST 21-MAR-2002
DEFINITION UI-CF-EC1-aca-j-15-0-UI-s1 UI-CF-EC1 Homo sapiens cDNA clone
ACCESSION BM974454
VERSION BM974454.1 GI:19592045
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 735)
Ronald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
PolyA-yes.
FEATURES
source
Location/Qualifiers
1..735
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-aca-j-15-0-UI"
/clone_1lb="UI-CF-EC1"
/tissue_type="lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Ronald, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(drr)18 tail. The sequence tag for this library is
AAGCGCTTAC
TAG_LIB=UI-CF-EC1
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_SEQ=AAGCGCTTAC"
BASE COUNT 236 a 139 c 163 g 197 t
ORIGIN
Query Match 30.8%; Score 717.4; DB 14; Length 735;
Best Local Similarity 99.9%; Pred. No. 8e-190;
Matches 718; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1610 TAGGCAATTCCTGCTGAGAGTTTCCCGAAGAGTACAGCCCGACTTCGACTCTACCT 1669
DB 734 TAGGCAATTCCTGCTGAGAGTTTCCCGAAGAGTACAGCCCGACTTCGACTCTACCT 675
QY 1670 GTACTATTGCAGACCTGTACACCTTCAGAGACTTGCACCCCGAGTTCCATGATTAAGTT 1729
DB 674 GTACTATTGCAGACCTGTACACCTTCAGAGACTTGCACCCCGAGTTCCATGATTAAGTT 615

QY 1730 ATCAAAAGATTATTCATCTCTCCCGTGATGAGAGATTGTGTGAATTTTCAAGCCTT 1789
DB 614 ATCAAAAGATTATTCATCTCTCCCGTGATGAGAGATTGTGTGAATTTTCAAGCCTT 555
QY 1790 CAGTTATTTCACCTATTTCGAAGAAGATAGATTAGTTGGCGGGCTCGAGCTATG 1849
DB 554 CAGTTATTTCACCTATTTCGAAGAAGATAGATTAGTTGGCGGGCTCGAGCTATG 495
QY 1850 TTCAAGAGCTGTGAACAGCTTGTGTCACCTCTTCACCTCTTCACCTCTTCGACCTG 1909
DB 494 TTCAAGAGCTGTGAACAGCTTGTGTCACCTCTTCACCTCTTCGACCTCTTCGACCTG 435
QY 1910 TGTACTGCTTTCCAAAGACCCGGAGCTGGCGGGAACCTGGAGTAGTACTTGGCT 1969
DB 434 TGTACTGCTTTCCAAAGACCCGGAGCTGGCGGGAACCTGGAGTAGTACTTGGCT 375
QY 1970 TTTTGGCTACAGAGAGAGCTATGTAAACAAACACAGAGATGCAAGGCTTTTGA 2029
DB 374 TTTTGGCTACAGAGAGAGCTATGTAAACAAACACAGAGATGCAAGGCTTTTGA 315
QY 2030 GAATGTGTTTCAAAACCATGCTGTATTTTCAACCATTAAGAGTTTCAGTTGCTT 2089
DB 314 GAATGTGTTTCAAAACCATGCTGTATTTTCAACCATTAAGAGTTTCAGTTGCTT 255
QY 2090 AATTTGTATACGGTTTAAATTTGCTGCTTGTATTTTGAATTTTAAATAATATGTC 2149
DB 254 AATTTGTATACGGTTTAAATTTGCTGCTTGTATTTTGAATTTTAAATAATATGTC 195
QY 2150 GTAGATTCCTTCGAAGGCTTCAGACACATGCTATGTCCTGCCCAACCCAGTC 2209
DB 194 GTAGATTCCTTCGAAGGCTTCAGACACATGCTATGTCCTGCCCAACCCAGTC 135
QY 2210 TCCTGTCAATTTTACGCCAGTGTGTTCTTTGAGAGACCCCTTAACTTGTCTTTAGAA 2269
DB 134 TCCTGTCAATTTTACGCCAGTGTGTTCTTTGAGAGACCCCTTAACTTGTCTTTAGAA 75
QY 2270 TTTTACCCCAATTTGATTTGGAATGAGAGTCCCAACATGATTAATTTTGAAGA 2328
DB 74 TTTTACCCCAATTTGATTTGGAATGAGAGTCCCAACATGATTAATTTTGAAGA 16
RESULT 14
BI837271 794 bp mRNA linear EST 04-OCT-2001
LOCUS BI837271
DEFINITION 6030905631 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5229519.5',
mRNA sequence.
ACCESSION BI837271
VERSION BI837271.1 GI:15948809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 794)
NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1577 row: n column: 16
High quality sequence stop: 781.
FEATURES
source
Location/Qualifiers
1..794
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

Db      547 TCAAAAAGATATATCAATGCTCCCTGATAGAAATGTTGCTGAATTTTCAAGCCCTTC 488
QY      1791 AGTTATTTTCCACTATTTTCAAGAAATATGATAGTTGGGGGCTGTAGTATAGT 1850
Db      487 ACTTATTTTCCACTATTTTCAAGAAATATGATAGTTGGGGGCTGTAGTATAGT 428
QY      1851 TCAAAAGAGCTGAAAGCTGCTGCTCCTTCTTCACTCTTCCACTCTCTCTCTCTCTCT 1910
Db      427 TCAAAAGAGCTGAAAGCTGCTGCTCCTTCTTCACTCTTCCACTCTCTCTCTCTCT 368
QY      1911 GTTACTGCTTTTGAAGAGACCCGGAGAGCTGGCGGGAACCTTGGAGTAGTATGTTGCTT 1970
Db      367 GTTACTGCTTTTGAAGAGACCCGGAGAGCTGGCGGGAACCTTGGAGTAGTATGTTGCTT 308
QY      1971 TTTGGCTACAGAGAGAGCTATGTAACAACACAGCAGCAGGATCGAAGGTTTAAAG 2030
Db      307 TTTGGCTACAGAGAGAGCTATGTAACAACACAGCAGCAGGATCGAAGGTTTAAAG 248
QY      2031 AATGCTTTCAAAACCATGCTGCTGCTTATTTCAACATATAAGAGTTTCAGTTGCTCTTA 2090
Db      247 AATGCTTTCAAAACCATGCTGCTGCTTATTTCAACATATAAGAGTTTCAGTTGCTCTTA 188
QY      2091 AATTTGTATACGGTTTAAATCTGCTCTTGTTCATTTTGAATTTTAAAAATATGTCG 2150
Db      187 AATTTGTATACGGTTTAAATCTGCTCTTGTTCATTTTGAATTTTAAAAATATGTCG 128
QY      2151 TAGAATCTCTTGAAGAGCTTCAAGACACATGCTATGCTGTCTTCCCAACACAGCT 2210
Db      127 TAGAATCTCTTGAAGAGCTTCAAGACACATGCTATGCTGTCTTCCCAACACAGCT 68
QY      2211 CCTCTCATTTTAAAGCCAGTGTCTTCTTGTAGAGACCCCTTATCTTGTCTTGTAGAT 2270
Db      67 CCTCTCATTTTAAAGCCAGTGTCTTCTTGTAGAGACCCCTTATCTTGTCTTGTAGAT 8
QY      2271 TTTTACC 2277
Db      7 TTTTACC 1

```

RESULT 12
 B1523975 721 bp mRNA linear EST 29-AUG-2001
 LOCUS 603052129F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5201767 5',
 DEFINITION mRNA sequence.
 B1523975
 ACCESSION B1523975.1 GI:15348767
 VERSION EST.
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 721)
 AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M1505 row: j column: 08
 High quality sequence stop: 716.
 Location/Qualifiers
 1..721
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5201767"
 /clone_11b="NIH_MGC_122"

```

/lab host="DH108"
/notes="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 026. Note:
this is a NIH-MGC Library."
BASE COUNT      155 a      233 c      175 g      158 t
ORIGIN
Query Match      30.9% Score 719.4; DB 13; Length 721;
Best Local Similarity 99.9% Pred. No. 2.2e-190;
Matches 720; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      3 CCGGGGCTCTCCCGGTGCTCTCTCAAGACTCGTCCGCCCTTGGAAATAAACACC 62
Db      1 CCGGGGCTCTCCCGGTGCTCTCTCAAGACTCGTCCGCCCTTGGAAATAAACACC 60
QY      63 GCGAGCCCGAGGGCCAGAGAGGCGAGCTGCCAGCTCTCCGGGGTCCGCCG 122
Db      61 GCGAGCCCGAGGGCCAGAGAGGCGAGCTGCCAGCTCTCCGGGGTCCGCCG 120
QY      123 CGAGCTTTCTTCGCGCTTGGCATCTCTCTCGCGGCTTGGCATGCGGAAATAA 182
Db      121 CGAGCTTTCTTCGCGCTTGGCATCTCTCTCGCGGCTTGGCATGCGGAAATAA 180
QY      183 AAGGATACCTACTGTTACATTTCTGCTCTCTTCCAGCCCTGGGAATGCAAGCC 242
Db      181 AAGGATACCTACTGTTACATTTCTGCTCTCTTCCAGCCCTGGGAATGCAAGCC 240
QY      243 ACACTGCAAGATGGCTTTTACCTGATCGGCACTGAGAGAGCTTTTATGATGATGA 302
Db      241 ACACTGCAAGATGGCTTTTACCTGATCGGCACTGAGAGAGCTTTTATGATGATGA 300
QY      303 ATGCCGAACCATCCCGAGAGGCTCCGAGAGACATGATGTTTAAACAAATGGCGG 362
Db      301 ATGCCGAACCATCCCGAGAGGCTCCGAGAGACATGATGTTTAAACAAATGGCGG 360
QY      363 GATTTATGATTTCCCGGACAAACCTGTATGAGAGGCGCTTACTGAAACCCCTACTC 422
Db      361 GATTTATGATTTCCCGGACAAACCTGTATGAGAGGCGCTTACTGAAACCCCTACTC 420
QY      423 GACCCCTACTCAGAGTCCGATCCAGAGAGCTGCCACACATCTCAGCTCCAAACTATCC 482
Db      421 GACCCCTACTCAGAGTCCGATCCAGAGAGCTGCCACACATCTCAGCTCCAAACTATCC 480
QY      483 CAGCATCTCAGAGGCTTTATATGCGCTTGGATACAGATGATGAAAGCAACCAATG 542
Db      481 CAGCATCTCAGAGGCTTTATATGCGCTTGGATACAGATGATGAAAGCAACCAATG 540
QY      543 TGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCAACCCACCAGATGTGAT 602
Db      541 TGTGATGTGAGAGAGTGTGCACAGATTTCCACAGTGCAACCCACCAGATGTGAT 600
QY      603 CAATACGAGAGGGGGTACACCTCTCTGACGAGAGAGATTTGGCTTTGGAAGGCCA 662
Db      601 CAATACGAGAGGGGGTACACCTCTCTGACGAGAGAGATTTGGCTTTGGAAGGCCA 660
QY      663 GTGCTTGAATGATTAATGCTGATGTTAGTCTGAGAGAGCTGTGTGGAATGTTCC 722
Db      661 GTGCTTGAATGATTAATGCTGATGTTAGTCTGAGAGAGCTGTGTGGAATGTTCC 720
QY      723 T 723
Db      721 T 721

```

RESULT 13
 BM974454/c

Db 721 TTCTGCTCTGCTCCCTCAGGCTTACATCCCTGCTGATGACACCGAGAGCTCCAGAAC 780
QY 1032 CATCAACGAATGTGACACAGAACACACAGTGCAC-CCTGACACAGACGTACAAAT 1090
Db 781 ATCAACCGAATGTGACACAGAACACACAGTGCAC-CCTGACACAGACGTACAAAT 840
QY 1091 T--ACAGAGGGGCTTCAATGATGACACCCATCCGCTGTGAGAG 1134
Db 841 TTACAAGGGGGCTTCAATGATGACACCCATCCGCTGTGAGAG 887

RESULT 9
B1764020 880 bp mRNA linear EST 25-SEP-2001
LOCUS 603043221F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183693 5',
DEFINITION mRNA sequence.
ACCESSION B1764020
VERSION B1764020.1 GI:15755598
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
AUTHORS 1 (bases 1 to 880)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1458 row: 1 column: 06
High quality sequence stop: 846.
Location/Qualifiers
1. 880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5183693"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site1: NotI; Site2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons; age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs; 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized by C.
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT 192 a 278 c 215 g 195 t
ORIGIN

Query Match 34.0%; Score 790.6; DB 13; Length 880;
Best Local Similarity 98.7%; Pred. No. 2.5e-210;
Matches 860; Conservative 0; Mismatches 4; Indels 7; Gaps 6;

QY 3 CCGGCGGCTCTCCCTGCTCTCCGACGACTGCGGCGCTGTGGAATAAACACCC 62
Db 11 CCGGCGGCTCTCCCTGCTCTCCGACGACTGCGGCGCTGTGGAATAAACACCC 70
QY 63 GCGAGCCCGGAGGCGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCG 122
Db 71 GCGAGCCCGGAGGCGCCAGAGAGAGCGAGCTGCCGAGCTCTCCGGGGGTCCGCCCG 130
QY 123 CGAGCTTTTCTGCGCTTGCATCTCTCTGCGGCGCTTGGACATGCCAGGAATAA 182

Db 131 CGAGCTTTTCTGCGCTTGCATCTCTCTGCGGCGCTTGGACATGCCAGGAATAA 190
QY 183 AAGGATACCTACTGTTACATTTGCGCTCTGCTTCCAGGCCCTGGGAATGCAGAGC 242
Db 191 AAGGATACCTACTGTTACATTTGCGCTCTGCTTCCAGGCCCTGGGAATGCAGAGC 250
QY 243 ACAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGAGAGTGTATGATATTGATGA 302
Db 251 ACAGTGCAGCAATGGCTTTGACCTGATGCGCAGTCAAGAGAGTGTATGATATTGATGA 310
QY 303 ATGCCGAGCAATCCCGAGGCTTCCGAGAGACATGATGTGTAAACCAAAATGGCG 362
Db 311 ATGCCGAGCAATCCCGAGGCTTCCGAGAGACATGATGTGTAAACCAAAATGGCG 370
QY 363 GTATTATGATTCCTCCGAGCAAAACCTGTATATGAGGCGCTTGTGAACCCCTATCC 422
Db 371 GTATTATGATTCCTCCGAGCAAAACCTGTATATGAGGCGCTTGTGAACCCCTATCC 430
QY 423 GACCCCTTACTAGGTCCTGACCCAGAGAGTGCACACCTGCTGAGCTCCAAATATCC 482
Db 431 GACCCCTTACTAGGTCCTGACCCAGAGAGTGCACACCTGCTGAGCTCCAAATATCC 490
QY 483 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAATGATGAAAGCAACCAATG 542
Db 491 CACGATCTCCAGGCTCTTATATGCGCTTGTGATACCAATGATGAAAGCAACCAATG 550
QY 543 TGTGATGTGAGAGTGTGCAACAGATTCACACAGTGCACACCCACCAATGCGAT 602
Db 551 TGTGATGTGAGAGTGTGCAACAGATTCACACAGTGCACACCCACCAATGCGAT 610
QY 603 CAATAGTGAAGGGGGGTACACCTGCTCCGACCGACGAGATTTGCTTGAAGGCCA 662
Db 611 CAATAGTGAAGGGGGGTACACCTGCTCCGACCGAGAGATTTGCTTGAAGGCCA 670
QY 663 GTGCTTAGACATTTGATATGTCCTGATGTTAGTACGACCA-GCTGTGTGGAATGTT 721
Db 671 GTGCTTAGACATTTGATATGTCCTGATGTTAGTACGACCA-GCTGTGTGGAATGTT 730
QY 722 CTGATCTCTATTTCTTGTGA--CATGCAACCCCTGTTTACCTCAATGAGATGAA-GGT 778
Db 731 CTGATCTCTATTTCTTGTGAATTTGCAATTCACACCCCTGTTTACCTCAATGAGATGAA-GGT 790
QY 779 CTGTCACAGA-TGTGAACAGAGTGTGCAACGAGAACCCCTGCTGCAAACTGCTCAAC 837
Db 791 CTGTCACAGAATGTGAGAGAGTGTGCAACGAGAACCCCTGCTGCAAA-CTGCTGCAAC 848
QY 838 ACCTAGGCTCTTTCATCTGCGCTGTGACC 868
Db 849 ACCTAGGATCTTTCACCTGCGCTGTGACC 879

RESULT 10
B0718885 835 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8241132 lupsk1_sympathetic_trunk Homo sapiens cDNA clone
DEFINITION IMAGE:6187063 5', mRNA sequence.
ACCESSION B0718885
VERSION B0718885.1 GI:21857782
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE NIH-MGC http://mhc.nci.nih.gov/
AUTHORS 1 (bases 1 to 835)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

QY	1247	CGGTGCCGGTGAATCTCTCCAAATGCAAGCCACGACCCGCTACCTCGGGGCTATTACA	1306
Db	301	CCGTTCCCGGTGAATCTTCCAAATGCAAGCCACGACCCGCTACCTCGGGGCTATTACA	360
QY	1307	TTTTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGTCGGCAACGGGCCCA	1366
Db	361	TTTTCCAGATCAAAATCTGGGAATGAGGGCAGAGAAATTTTACATGTCGGCAACGGGCCCA	420
QY	1367	TCAGTGGCACCCCTGGTGATGACACGCCCCCATCAAGGGCCCCGGGAAATCCAGCTGACT	1428
Db	421	TCAGTGGCACCCCTGGTGATGACACGCCCCCATCAAGGGCCCCGGGAAATCCAGCTGACT	480
QY	1427	TGGAAATGATCAGCTGTCAACACTGTCAATCACTTACAGAGGCAAGCTCCGTATCCGACTGC	1488
Db	481	TGGAAATGATCAGCTGTCAACACTGTCAATCACTTACAGAGGCAAGCTCCGTATCCGACTGC	540
QY	1487	GGATATATGTGTGCGCAGTACCATTCTGTAGAGCTCGGGCTGAGACCTCCGACGCTCTC	1546
Db	541	GGATATATGTGTGCGCAGTACCATTCTGTAGAGCTCGGGCTGAGACCTCCGACGCTCTC	600
QY	1547	CATTGCGACCAAGGAGACAGGAAGAAGAAATTAACAGAGA-GAATGAGAGCGACACAG	1605
Db	601	CATTGCGACCAAGGAGACAGGAAGAAGAAATTAACAGAGAGCAATGAGAGCGACACAG	660
QY	1606	ACGTTAGGCAATTTCCCTGTGAAGCTTTCCCGAAGAGTGAAGCCGACATTCGAGACTC	1665
Db	661	ACGTTAGGCAATTTCCCTGTGTGAAGCTTTCCCGAAGAGTGAAGCCGACATTCGAGACTC	720
QY	1666	ACCTGTACTATTTGCGAGACTGTGCACCCGACAGAGACTTGGCACCCCA-GTTCTATGATA	1724
Db	721	AACTGTACTATTTGCGAGACTGTGCACCCGACAGAGACTTGGCACCCCAAGTTCTATGATA	780
QY	1725	CAGTTATCAAAAAGTATTATCATTTGCTCCCTGATAGAAAGATTGTTGGTAATTTTCAAG	1784
Db	781	CAGTTATCAAAAAGTATTATCATTTGCTCCCTGATAGAAAGATTGTTGGTAATTTTCAAG	839
QY	1785	GCCCTCAGTTATTTTCCACTATTTTCAAGAAATATAGATTAGTTGGCGG	1835
Db	840	GCCTTGG--CAATTCCTCTATTTCACAGGAACATGATAGATGACAGGG	888

	/clone="IMAGE:6187340"
	/clone_lib="lupskl_sym pathetic_trunk"
	/sex="male"
	/tissue_type="sympathetic trunk"
	/dev_stage="adult, 16 yr"
	/lab_host="DH10B"
	/note="vector: PCMV-SPORT6 (Life Technologies); Site_1:
	Note; Site_2: SalI; cDNA made by oligo-dT priming.
	directionally cloned using the following adaptors:
	5'-TGACCCAGGCGTCGG-3' and
	5'-GCATGTCTTAGATCGCAGGCGGCCGCT(15)-3'. Size selected >
	1 kb for average insert length 1.9 kb. This is a primary
	library, non-amplified. Library constructed by Life
	Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
	College of Medicine), available through Life
	Technologies."
BASE COUNT	219 a 264 c 219 g 202 t
ORIGIN	
Query Match	34.0%; Score 792.2; DB 14; Length 904;
Best Local Similarity	97.1%; Pred. No. 8.e-211;
Matches	Conservative 861; Mismatches 18; Indels 8; Gaps 5
Dy	256 GCCTTTACCTTGATCCCAGTCAGGACAGTGTATTAGTAATGAAATGCCGAACAATC 315
Dd	1 GGCTTTGACCCTGGATCCGCACTCAGAACAGTGTATTGATAAATGGCGGCTATTATCATTC 60
OY	316 CCCGAGGCGCTCCGAGGAGACATGATGTGTCTTAACCAAATGGCGGCTATTATGACATT 375
Dd	61 CCGGAGGCGCTCCGAGGAGACATGATGTGTCTTAACCAAATGGCGGCTATTATGACATT 120
OY	376 CCCCCGACAACCTGTGTATCGAGGCGCTACTTCGAACCCCTACTCGACCCCTCATCA 435
Dd	121 CCCCCGACAACCTGTGTATCGAGGCGCTACTTCGAACCCCTACTCGACCCCTCATCA 180
OY	436 GGTCGGTACCCAGAGCTGCCCAACATCTCAGTCCAAACTATCCAGCATGCCAGG 495
Dd	181 GGTCGGTACCCAGAGCTGCCCAACATCTCAGTCCAAACTATCCAGCATGCCAGG 240
OY	496 CCTCTTATATGCGCCTTTGGATTACCAATGATGAAGAACAACAATGTGTGATGTGAC 555
Dd	241 CCTCTTATATGCGCCTTTGGATTACCAATGATGAAGAACAACAATGTGTGATGTGAC 300
OY	556 GAGGTGGACAAGATGCCACAGTGCACCCACCACCGATCTGCATCATATCGAAGGC 615
Dd	301 GAGGTGGACAAGATGCCACAGTGCACCCACCACCGATCTGCATCATATCGAAGGC 360
OY	616 GGGTACACCTTCTCTGTCACGAGCGAGATATGGCTTCGGAAGGCAGTGTATGACATT 675
Dd	361 GGGTACACCTTCTCTGTCACGAGCGAGATATGGCTTCGGAAGGCAGTGTATGACATT 420
OY	676 GATGAATGTGCTATGGTTACTGCCACAGCTGTGTGCGAATGTTTCTGTGATCTATTC 735
Dd	421 GATGAATGTGCTATGGTTACTGCCACAGCTGTGTGCGAATGTTTCTGTGATCTATTC 480
OY	736 TGTACATGCAACCTCGTTTACCTCCTCATAGGATGGAAGCTCTTGCCAAGATGTGAC 795
Dd	481 TGTACATGCAACCTCGTTTACCTCCTCATAGGATGGAAGCTCTTGCCAAGATGTGAC 540
OY	796 GAGTGTGCCACCGAAGAACCCCTGCGTCAACCTGTGCTCAACACACTAGCGCTCTTATC 855
Dd	541 GAGTGTGCCACCGAAGAACCCCTGCGTCAACCTGTGCTCAACACACTAGCGCTCTTATC 600
OY	856 TTGCGCGTGTGCCAGAGATATGAACCTTGAGGAATGGCTTCATTGTAGAGATATGAC 915
Dd	601 TTGCGCGTGTGCCAGAGATATGAACCTTGAGGAATGGCTTCATTGTAGAGATATGAC 660
OY	916 GAGTCAAGCTTCTCTGAGTCTCTTGCCCAACATGAGTGTGAACACGCCGCGACATAC 975
Dd	661 GAGTCAAGCTTCTCTGAGTCTCTTGCCCAACATGAGTGTGAACACGCCGCGACATAC 720
OY	976 TTCGTCTGCTGCGCTCCAGGCTACAT--CTGCTGATATCACAC--GAACTGTCCAGA-- 1031

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 966)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@rs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM13583 row: f column: 05
 High quality sequence stop: 634.

FEATURES

source

1. 996
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6188164"
 /clone_1lb="Lupski_sympathetic_trunk"
 /sex="male"
 /issue_type="sympathetic trunk"
 /dev_stage="adult, 16 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors:
 5'-TCGACCCGACGCTCCG-3' and
 5'-GACTACTGTAGATCGAGCGAGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies."

BASE COUNT 245 a 284 c 242 g 224 t 1 others
 ORIGIN

Query Match 35.5%; Score 825.4; DB 14; Length 996;
 Best Local Similarity 97.9%; Pred. No. 4,4e-220;

Matches 889; Conservative 0; Mismatches 12; Indels 7; Gaps 5;

219 TCCAGCCCTGGGAATGACAGCAGCAGTGCAGATGGCTTACCTGATGCGCAGTC 278
 1 TTCCAGCCCTGGGAATGACAGCAGCAGTGCAGATGGCTTACCTGATGCGCAGTC 60
 279 AGGACAGTGTAGATATTGATGATGCGGACCATCCCGAGGCTGCCGAGAGACAT 338
 61 AGGACAGTGTAGATATTGATGATGCGGACCATCCCGAGGCTGCCGAGAGACAT 120
 339 GATGTGTGTTAACCAATATGGCGGTATTATGATTCCTCCCGACCAACCTGTGTATCG 398
 121 GATGTGTGTTAACCAATATGGCGGTATTATGATTCCTCCCGACCAACCTGTGTATCG 180
 399 AGGCGCTTACTGGAACCCCTACTGACCCCTACTAGTCCGTAACCCAGAGCTGCC 458
 181 AGGCGCTTACTGGAACCCCTACTGACCCCTACTAGTCCGTAACCCAGAGCTGCC 240
 459 ACCACTCTCAGCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTTGATA 518
 241 ACCACTCTCAGCTCCAACTATCCACGATCTCCAGGCTCTTATATGCGCTTTGATA 300
 519 CCAGTGTGATGAAGAACCAATATGTTGATGTGGAGCATGTGCAACAGATTCCACCA 578
 301 CCAGTGTGATGAAGAACCAATATGTTGATGTGGAGCATGTGCAACAGATTCCACCA 360
 579 GTGCAACCCACCCAGATCTGCATCAATCTGAAAGGGGGTACACGCTGCTCCAGACGA 638
 361 GTGCAACCCACCCAGATCTGCATCAATCTGAAAGGGGGTACACGCTGCTCCAGACGA 420
 639 CGGATATTGGCTTCTGGAAGGCCAGTGTGATGACATTGATGATGCTATGCTTACTG 698

|||||
 Db 421 CGGATATTGGCTTCTGGAAGGCCAGTGTGATGACATTGATGATGCTGATGTTACTG 480
 Qy 699 CCAGCAGCTGTGCGAATGTTCTCGATCCTTATTTGTATCATCAACCTGGTTTAC 758
 Db 481 CCGAGAGCTGTGCGAATGTTCTCGATCCTTATTTGTATCATCAACCTGGTTTAC 540
 Qy 759 CCATATGAGATGGAAGGCTTGTCCAAAGATGTGAAGAGTGTCCACGAGAACCCCTG 818
 Db 541 CCTCAATGAGATGGAAGGCTTGTCCAAAGATGTGAAGAGTGTCCACGAGAACCCCTG 600
 Qy 819 CGTCAACCTGGGTCAACACCTACGAGCTTTTATCTGCGCTGACCCAGATATGA 878
 Db 601 CGTCAACCTGGGTCAACACCTACGAGCTTTTATCTGCGCTGACCCAGATATGA 660
 Qy 879 ACTTGAGGAAGATGCGCTTATTCATGATGATATGAGCAGATGACCTTCTGACTTCT 938
 Db 661 ACTTGAGGAAGATGCGCTTATTCATGATGATATGAGCAGATGACCTTCTGACTTCT 720
 Qy 939 CTGCAACATGAGTGTGTAACCGACCGGACATCTTGTCTCTGCTGCC-TCGAGCT 997
 Db 721 CTGCAACATGAGTGTGTAACCGACCGGACATCTTGTCTCTGCTGCC-TCGAGCT 780
 Qy 998 ACATCTGTGATGACACCGAAGCTGCAAGACATCAAGATGTGACAGCAGG-AAAC 1056
 Db 781 ACATCTGTGATGACACCGAAGCTGCAAGACATCAAGATGTGACAGCAGG-AAAC 840
 Qy 1057 CACACGT-GCAACCTGACGAG-ACGTCTCAATTT--ACAGGGGCTTCAATGCA 1111
 Db 841 CACACGTGGAACCTGACGAGACGTCCTCAATTTACCAAGGGGCTTCAATGCA 900
 Qy 1112 TCGACCC 1119
 Db 901 TCGACCC 908

RESULT 6
 LOCUS B1825121 882 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603038662F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179550 5',
 mRNA sequence.
 ACCESSION B1825121
 VERSION B1825121.1 GI:15936671
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 882)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@rs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM11447 row: 1 column: 15
 High quality sequence start: 14
 High quality sequence stop: 844.

FEATURES

source

1. 882
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5179550"
 /clone_1lb="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA

	Matches	849;	Conservative	1;	Mismatches	26;	Indels	1;	Gaps	1;
QY	1416	CCACGTGGACTTGGAAATGATCACTGTCAACACACTGTCAATCACTTCAGAGCGACCTCGT	1475							
Db	876	CCACGTGGACTTGGAAATGATCACTGTCAACACACTGTCAATCACTTCAGAGCGACACNCCTG	817							
QY	1476	GATCCGACTCGGGATATATGTGTGGAGTACCCATTCCTGAGCCCTCGGGCTGGAGCCCTCGG	1535							
Db	816	GATCCGACNCGGAGNANANGTNCAGTACCCANNCTGAGCCCTC-GGCTTGGAGCCNCCG	758							
QY	1536	ACGCTGCCTCTCATTTGGCACCAAGGAGCAGAGAGAAGAGAGAAATTAACAGAGAAATGAG	1595							
Db	757	ACGCTGCCTCTCATTTGGCACCAAGGAGCAGAGAGAAGAGAAANNACAGAGAAATGAG	698							
QY	1596	AGGACACACAACGTTAGGCATTTCTCTGTAACCTTTCCCGAAGAGTCACCCCCGACTT	1655							
Db	697	AGGACACACAACGTTAGGCATTTCTCTGTAACCTTTCCCGAAGAGACACCCCCGACANN	638							
QY	1656	CCGTACCTCAACCTGTACTATTGGAGACCTGTCAACCTGACAGACTGCCACCCCCAGTT	1715							
Db	637	CCGTACCTCAACCTGTACTATTGGAGACCTGTCAACCTGACAGACTTGCACCCCCAGTT	578							
QY	1716	CCATGATACAGTTATCAAAAAAGTATTATCAATTCCTCCCTGATAGAGAAGATTGTTGTGA	1775							
Db	577	CCATGACACAGTTATCAAAAAAGTATTATCAATTCNCCNCCCTGATAGAGAAGATTGTTGTGA	518							
QY	1776	ATTTTCAAGGCTTCAGTTTATTTCCACATATTTCCAAAAGAAAATAGATTAGTTTGGCGG	1835							
Db	517	ATTTTCAAGGCTTCAGTTTATTTCCACATATTTCCAAAAGAAAATAGATTAGTTTGGCGG	458							

Db 457 GGTCTGAGCTGTATGTTCAAGACGTGTGACCTCTTCTTCACCTCTTCCAC 398

QY 1896 TCCCTTCTCTCACTGTGTACTGCTTTTGCAAGACCCGGGAGCTTGGCGGGCAACCTTGGGA 1955

Db	397	TCCCTTCCTCACTGCTGTACGCTTCCCAAGACCCGGGGGGCTGGCGGGAGACCTGGGA	338
QY	1956	GTACACTAGTTTGGCTTTTGGCGTACACAGAGACGGCTATGTAAACAAACACAGCAGATC	2015
Db	337	GTACACGTTTGGCTTTTGGCGTACACAGAGAGGCTATGTAAACAAACACAGAGGATC	278

2016 GAAGGTTTTCAGATGTCTTTCAAACCATGCGTATTTTCAACCATAAAGAAG 2075

Dd	277	GAAGGGTTTTAGAGAAATGTGTTCCAAAACCATGCCTGGATTTTCAACCATAAAGAAG	218
Dy	2076	TTTCAGTTGTCCTTAAATTTGTADAAAGCGTTTAATTCGTCTTGTTCAATTTGAGATATT	21355

2136 TTTAAATAATGTCGTGAAATTCCTTGGAAGCCCTTCACACACATGCTGATGTTCTGTCT 21395

Db 157 TTAAAAATATGTCGTAGATTCTTCGAAAGGCGCTCAGACACATGCTATGTTCTGTCT 98

QY	2196	TCCCAAAACCCAGTGCCTCCATTTAGCCCAATGTTTTCTTTGAGAGACCCCTTAATCT	2255
Db	97	TCCCAAAACCCAGNCTCCCTCCATTTAGCCCAATGTTTTCTTTGAGAGACCCCTTAATCT	38
QY	2256	TGCTTTCTTTAGAAATTTTACCCAATGGATTTGAAT	2292
Db	37	TGCTTTCTTTAGAAATTTTACCCAATGGATTTGAAT	1
RESULT 5			
BQ720516			
LOCUS	BQ720516	996 bp	mRNA
DEFINITION	AEENOCOURT_8219503	lupsk1_symptomatic_trunk	Homo sapiens cDNA clone
	IMAGE:6188164	5'	mRNA sequence.

ACCESSION	BO720516	GI:21859413
VERSION	BO720516.1	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

OY	210	TTCTGTTTCCAAAGCCCTGGGAATGCAACAGGACAGTGCACGATGGCTTTGACCTGGA	269
Db	129	TTCTGTTTCCAAAGCCCTGGGAATGCAACAGGACAGTGCACGATGGCTTTGACCTGGA	188
OY	270	TCGCGAGTCAGAGCAGTGTATTGATATGATGAATGCGGAACCAATCCCGAGGCTGCGG	329
Db	189	TCGCGAGTCAGAGCAGTGTATTGATATGATGAATGCGGAACCAATCCCGAGGCTGCGG	248
OY	330	AGGAGACATGATGTGTGTAAACCAAAATGGCGGGATTTATATGATTTCCCGGACAACCC	389
Db	249	AGGAGACATGATGTGTGTAAACCAAAATGGCGGGATTTATGCAATTTCCCGGACAACCC	308
OY	390	TGTATATGAGAGGCCCTACTCGAACCCCTACTCGAACCCCTACTCGAACCCCTACTCGAAC	449
Db	309	TGTATATGAGAGGCCCTACTCGAACCCCTACTCGAACCCCTACTCGAACCCCTACTCGAAC	368
OY	450	AGCTGCCCCACCACTCTCAGCTCCAAACTATCCACGATCTCCAGGCCCTCTTATATGCCG	509
Db	369	AGCTGCCCCACCACTCTCAGCTCCAAACTATCCACGATCTCCAGGCCCTCTTATATGCCG	428
OY	510	CTTTGGATACCAAGATGATGAAAGCAACCAATGTGTGATGTGACAGAGTGTGCACACGA	569
Db	429	CTTTGGATACCAAGATGATGAAAGCAACCAATGTGTGATGTGACAGAGTGTGCACACGA	488
OY	570	TTTCCACAGTGCACAAACCCACCCAGATGTGATCAATCTGAAAGCGGGGTACACTGTGTC	629
Db	489	TTTCCACAGTGCACAAACCCACCCAGATGTGATCAATCTGAAAGCGGGGTACACTGTGTC	548
OY	630	CTGCACCGACGGAATATTTGGCTTCTGGAAGGCCAGTGTCTTAGACATTTGATGTGCTA	689
Db	549	CTGCACCGACGGAATATTTGGCTTCTGGAAGGCCAGTGTCTTAGACATTTGATGTGCTA	608
OY	690	TGTTTACCTGCGCAGAGCTGTGGGAATGTTCTCGAGATCCATTTCTGTACATGCAACCC	749
Db	609	TGTTTACCTGCGCAGAGCTGTGGGAATGTTCTCGAGATCCATTTCTGTACATGCAACCC	668
OY	750	TGTTTTTACCCTCATGAGATGAGATGGAAGTCTTGGCCAAAGATGTGAACAGAGTGTGCCACGA	809
Db	669	TGTTTTTACCCTCATGAGATGAGATGGAAGTCTTGGCCAAAGATGTGAACAGAGTGTGCCACGA	728
OY	810	GAACCCCTGCGTGCACAACTGCTGCACACCTACGAGCTCTTTCATCTGCGCTGTGACCC	869
Db	729	GAACCCCTGCGTGCACAACTGCTGCACACCTACGAGCTCTTTCATCTGCGCTGTGACCC	788
OY	870	AGGATATGAATTGAGGAAGATGGCGTTCATTTGAGTGAATATGAGAGAGTGCAGCTTCTC	929
Db	789	AGGATATGAATTGAGGAAGATGGCGTTCATTTGAGTGAATATGAGAGAGTGCAGCTTCTC	848
OY	930	TGAGTTCTCTGCGCAACATGTAGTGTGTGACACAGCCCGGCACATATCTCTCTCTGCC	989
Db	849	TGAGTTCTCTGCGCAACATGTAGTGTGTGACACAGCCCGGCACATATCTCTCTCTGCC	908
OY	990	TCCAGGCTACATCTCTGCTGATGACAAACCGAAGCTGCCAGACATCAACGAATGTGACCA	1049
Db	909	T-CAGGCTACATCTCTGCTGATGACAAACCGAAGCTGCCAGACATCAACGAATGTGACCA	966
OY	1050	CAGAACCACACGCTGCAACCTGCGACGAGAGCTGTACAAATTACAGGGGG	1100
Db	967	CAGAACCACACCTGCAACCTGCGACGAGAGCTGTACAAATTACAGGGGG	1014
RESULT 3			
BM545444			
LOCUS			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
EukeyrOta: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT		
1 (bases 1 to 954)	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	NIH-MGC http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.	Email: cgabbs-remail.nih.gov	Tissue Procurement: Invitrogen			
	CDNA Library Preparation: Life Technologies, Inc.	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	DNA Sequencing by: Agencourt Bioscience Corporation	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov		
	Plate: LLM12727	row: c	column: 05	High quality sequence stop: 676.		
FEATURES	source	location/Qualifiers				
	1..954	/organism="Homo sapiens"				
	/db_xref="taxon:9606"					
	/clone IMAGE:5729596"					
	/clone_lib="NIH_MGC_124"					
	/tissue_type="hippocampus"					
	/lab_host="DH10B"					
	/note="Organ: brain; Vector: pCMV-Sport6; Site_1: EcoRV (destroyed); Site_2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."					
BASE COUNT	232 a	278 c	238 g	206 t		
ORIGIN						
Query Match	39.8%	Score 925.4	DB 13	Length 954		
Best Local Similarity	99.3%	Pred. No. 3.8e-248				
Matches 940	Conservative 0	Mismatches 6	Indels 1	Gaps 1		
QY	742	TGCAACCTGGTTTATCCCTCAATGAGAGTGAAGGCTTCCCAAGATGTAGACAGTGT	801			
DB	8	TGCAACCTGGTTTATCCCTCAATGAGAGTGAAGGCTTCCCAAGATGTAGACAGTGT	67			
QY	802	GCCACCGAGAACCCCTGCGTCAAAACCTGCGTCAACACCTACGCGCTTTTCATCTGCCG	861			
DB	68	GCCACCGAGAACCCCTGCGTCAAAACCTGCGTCAACACCTACGCGCTTTTCATCTGCCG	127			
QY	862	TGTGACCCAGATATGAACTTGAGAGAAATGGCGTTTCAATTGCAGATATGACAGAGTC	921			
DB	128	TGTGACCCAGATATGAACTTGAGAGAAATGGCGTTTCAATTGCAGATATGACAGAGTC	187			
QY	922	AGCTTCTGTGAGTTCCTCTGCGCAACATGATGTGAAACAGCCCGACATCTTCTGCG	981			
DB	188	AGCTTCTGTGAGTTCCTCTGCGCAACATGATGTGAAACAGCCCGACATCTTCTGCG	247			
QY	982	TGCTGCCCTCAGGCTACATCTCTGCGATGACAAACGGAAGCTGCCAAGACATCAAGAA	1041			
DB	248	TGCTGCCCTCAGGCTACATCTCTGCGATGACAAACGGAAGCTGCCAAGACATCAAGAA	307			
QY	1042	TGTAGACACAGAACACACAGTGCACCTGTGACAGACAGCTGTACATTTCAAGGGGCG	1101			
DB	308	TGTAGACACAGAACACACAGTGCACCTGTGACAGACAGCTGTACATTTCAAGGGGCG	367			
QY	1102	TTTCAAAAGCATCGACCCCATCCGCGTGTGAGAGGCTTATCTGAGAGATCAGATTAACGC	1161			
DB	368	TTTCAAAAGCATCGACCCCATCCGCGTGTGAGAGGCTTATCTGAGAGATCAGATTAACGC	427			
QY	1162	TGTATGTGTCTGCTGTGAAACCTCTGGCGTGCAGAGACCAAGCCCTTTACCATTTGTACCG	1221			
DB	428	TGTATGTGTCTGCTGTGAAACCTCTGGCGTGCAGAGACCAAGCCCTTTACCATTTGTACCG	487			
QY	1222	GACATGACGTGTGTGAGACGCTCCGCTTCGCGTGCATCTTCCAAATGCAAGCCACG	1281			

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 293 a 224 c 262 g 268 t 6 others

Query Match 44.8%; Score 1042.2; DB 9; Length 1053;
Best Local Similarity 99.0%; Pred. No. 6.6e-281;
Matches 1042; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

QY 1242 ACCGTCGCTCCCGCTGACATCTCCAAATGCAACGCCAGCCCTACCTGGGCGCTA 1301
DB 1053 ACCGTCGCTCCCGCTGACATCTCCAAATGCAACGCCAGCCCTACCTGGGCGCTA 994
QY 1302 TTACATTTTCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAAAAGG 1361
DB 993 TTACATTTTCAGATCAAAATCTGGGAATGAGGCGAGAAATTTTACATCGGCAAAAGG 934
QY 1362 CCCCATCAGTCCACCTGGTGTATGACACGCCCATCAAAAGGCCCCGGGAAATCCAGCT 1421
DB 933 CCCCATCAGTCCACCTGGTGTATGACACGCCCATCAAAAGGCCCCGGGAAATCCAGCT 874
QY 1422 GGACTTGGAAATGATCACTGTCACACATGTCATCAATCTGAGAGCAGCTCGGTATCCG 1481
DB 873 GGACTTGGAAATGATCACTGTCACACATGTCATCAATCTGAGAGCAGCTCGGTATCCG 814
QY 1482 ACTGCGGATATATGTGTGCGAGTACCCATTCGTGAGCCCTGGGCTCGAGCGCTG 1541
DB 813 ACTGCGGATATATGTGTGCGAGTACCCATTCGTGAGCCCTGGGCTCGAGCGCTG 754
QY 1542 CCTCTCTATTGGCACAAGGAGCAGAGAGAGAAATTAACAGAGAGATGAGAGCGAC 1601
DB 753 CCTCTCTATTGGCACAAGGAGCAGAGAGAGAAATTAACAGAGAGATGAGAGCGAC 694
QY 1602 ACAGAGCTTGAAGCTTCTGCTGACAGCTTCCCGGAAAGTCAAGAGTCCGCTGAC 1661
DB 693 ACAGAGCTTGAAGCTTCTGCTGACAGCTTCCCGGAAAGTCAAGAGTCCGCTGAC 634
QY 1662 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTGCACCCCGAGTCTCATG 1721
DB 633 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTGCACCCCGAGTCTCATG 574
QY 1722 ATACAGTTATCAAAATATTTATCATGCTCCCTGATAGAAATTTGTTGTAATTTTC 1781
DB 573 ATACAGTTATCAAAATATTTATCATGACCCCGATAGAAATTTGTTGTAATTTTC 514
QY 1782 AAGGCTTCAGTTATTTCCACATTTTCAAGAAATATAGATTAGTTGGCGGGCTCG 1841
DB 513 AAGGCTTCAGTTATTTCCACATTTTCAAGAAATATAGATTAGTTGGCGGGCTCG 454
QY 1842 AGCTATATGTCAAAGACTGTGAACAGCTTGTCTACCTCTTCACCTCTCTTC 1901
DB 453 AGCTATATGTCAAAGACTGTGAACAGCTTGTCTACCTCTTCACCTCTCTTC 394
QY 1902 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTGCACCCCGAGTCTCATG 1961
DB 393 TCTCAGCTGATATTCAGACCTGTCAACCTGACAGACTTGCACCCCGAGTCTCATG 334
QY 1962 AGTTGCTTTTTCAGACAGAGAGGCTATGTAAACAAACACAGCAGATCGAAGG 2021
DB 333 AGTTGCTTTTTCAGACAGAGAGGCTATGTAAACAAACACAGCAGATCGAAGG 274
QY 2022 TTTTATGAGAAATGTTTCAAAACCATGCTGTATTTTCAACCATATAAAGAGTTTCAG 2081
DB 273 TTTTATGAGAAATGTTTCAAAACCATGCTGTATTTTCAACCATATAAAGAGTTTCAG 214
QY 2082 TTGTCTTAATTTGTTAAGGTTTAAATCTGTCTGTCTCATTTTGTGATTTTAA 2141

```

```

DB 213 TTGTCTTAATTTGTTAAGGTTTAAATCTGTCTGTCTCATTTTGTGATTTTAA 1154
QY 2142 AATATGCTGTAATTCCTTGAAGGCTTCAGACATGCTATGCTGCTCCCA 2201
DB 153 AATATGCTGTAATTCCTTGAAGGCTTCAGACATGCTATGCTGCTCCCA 94
QY 2202 ACCAGTCTCTCTCCATTTTACCCAGTGTCTTTGTTGAGAGACCCCTTAATCTGCTTT 2261
DB 93 ACCAGTCTCTCTCCATTTTACCCAGTGTCTTTGTTGAGAGACCCCTTAATCTGCTTT 34
QY 2262 CTTTGAATTTTAAACCAATGATGGAATGC 2294
DB 33 CTTTGAATTTTAAACCAATGATGGAATGC 1

```

RESULT 2
BM921371 1014 bp mRNA linear EST 13-MAR-2002
LOCUS
DEFINITION
AGENCOURT.6626255 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752785
5', mRNA sequence.
ACCESSION
BM921371
VERSION
BM921371.1 GI:19371750
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1014)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
plate: LLM1287 row: 1 column: 10
High quality sequence stop: 738.

FEATURES
source
Location/Qualifiers
1..1014

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752785"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA
source: anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research genetics tracking code
021. Note: this is a NIH-MGC Library."

BASE COUNT 235 a 300 c 246 g 233 t

ORIGIN

Query Match 40.5%; Score 943.4; DB 14; Length 1014;
Best Local Similarity 98.4%; Pred. No. 3.4e-253;
Matches 995; Conservative 0; Mismatches 11; Indels 5; Gaps 4;

```

QY 90 GAGCTGCCGAGCTCTCCGGGGGTCGCCCGGAGCTTTCTCGCCTTCGATC 149
DB 9 GATGTGCCCGAGCTCTCCGGGGGTCGCCCGGAGCTTTCTCGCCTTCGATC 68
QY 150 CTCTCGCGGCTTGGACATGCCAGAAATAAAGATACACGTTACATCTGCG 209
DB 69 CTCTCGCGGCTTGGACATGCCAGAAATAAAGATACACGTTACATCTGCG 128

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:48:06 ; Search time 3433.81 Seconds
(Without alignments)
10979.953 Million cell updates/sec

Title: US-09-674-379A-12
Perfect score: 2328
Sequence: 1 gaccgcgcgcctcccccgtg.....tgattaataattgaagaga 2328

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1042.2	44.8	1053 9	AL571218 AL571218
2	943.4	40.5	1014 14	BM921371 AGENCOURT
3	925.4	39.8	954 13	BM545444 AGENCOURT
4	837.4	36.0	876 9	AL572040 AL572040
5	825.4	35.5	996 14	BM720516 AGENCOURT
6	807.6	34.7	882 13	BM825121 BM825121

Result No.	Score	Query Match length	ID	Description
7	793	34.1	912 13	BM768516
8	792.2	34.0	904 14	BM716569
9	790.6	34.0	880 14	BM764020
10	785	33.7	835 14	BM718885
11	723.8	31.1	727 9	AL570066
12	719.4	30.9	721 13	BM523975
13	717.4	30.8	735 14	BM974454
14	700.4	30.1	794 13	BM837271
15	700.4	30.1	1006 9	AL545549
16	682.4	29.7	709 14	BM015079
17	683.4	29.4	699 14	BM970551
18	674.6	29.0	762 13	BM834712
19	666.8	28.6	891 12	BM339448
20	665	28.6	822 12	BM109391
21	650.4	27.9	667 12	BM970136
22	650.4	27.9	669 13	BM666348
23	646.4	27.8	889 14	BM897144
24	639.2	27.5	796 13	BM453348
25	637.6	27.4	870 14	BM930927
26	636	27.3	644 13	BM938621
27	630.4	27.1	672 14	BM006021
28	626	26.9	803 12	BM786513
29	620.6	26.7	685 9	AL543659
30	615.2	26.4	1015 13	BM544363
31	608.4	26.1	627 14	BM772930
32	606	26.0	629 13	BM489863
33	600.2	25.8	906 14	BM723059
34	592.8	25.5	717 13	BM251219
35	584.6	25.1	886 12	BM485593
36	576.8	24.8	690 12	BM104873
37	576.4	24.8	933 14	BM716088
38	575.8	24.7	770 12	BM542457
39	573.6	24.6	627 12	BM909699
40	573	24.6	573 10	BM183044
41	571	24.5	587 14	BM690459
42	570.2	24.5	586 14	BM695168
43	568	24.4	586 13	BM938622
44	565.6	24.3	572 10	BM299380
45	564.8	24.3	815 12	BM174360

ALIGNMENTS

RESULT 1
LOCUS AL571218/c
DEFINITION AL571218 LTI_NFL006.Pl2 Homo sapiens CDNA clone CS0D1015YE20 3
ACCESSION AL571218
VERSION AL571218.1 GI:12928294
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1053)
AUTHORS Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1015YE20"
/clone_lib="LTI_NFL006.Pl2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand CDNA was primed with a NotI-oligo(dT) primer. Five prime end

XX (ONOX) ONO PHARM CO LTD.
PA
XX

PI Honjo T, Tashiro K, Nakamura T;
XX

DR WPI: 2000-038646/03.
XX

DR N-PSDB; AA239385.
XX

PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX

PS Claim 1; Page 51-53; 70pp; Japanese.
XX

CC The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/repatriation
CC activity, actin/inhibin activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadherin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents the mature protein of the invention which
CC can be used for modulating smooth muscle cell proliferation.
XX

SQ Sequence 423 AA:

Query Match 90.9%; Score 2302; DB 21; Length 423;
Best Local Similarity 94.8%; Pred. No. 1.4e-148;

Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 26 OCTNGFDLDROSGQCDDIDECRTIPACRGDMCVNONGYLCPRTNPYRGPSNPYS 85
DB 1 OCTNGFDLDROSGQCDDIDECRTIPACRGDMCVNONGYLCPRTNPYRGPSNPYS 60
QY 86 TPYSGPYPAAPPLASAPNTPTISPLICRPGYQMDSESNQCVDDECATDSHCNPTQICI 145
DB 61 TSYSGPYPAAPPLASAPNTPTISPLICRPGYQMDSESNQCVDDECATDSHCNPTQICI 120
QY 146 NTEGGYTCSTDGWLLBEGQCDDIDECRYGYCOQLCANVPQSYSCGPNPGFTLNEDGRSC 205
DB 121 NTEGGYTCSTDGWLLBEGQCDDIDECRYGYCOQLCANVPQSYSCGPNPGFTLNEDGRSC 180
QY 206 ODVNECATENPCVQTCVNTYSGFICRCDPGYLEEDGVHCSMDCESEFLCOHECVNQ 265
DB 181 ODVNECATENPCVQTCVNTYSGFICRCDPGYLEEDGVHCSMDCESEFLCOHECVNQ 240
QY 266 PGTYFCSCPPIYLDDNRSQCDDINECEHNHTCNLQOTCYNLOGGFRCIDPINCCEPYL 325
DB 241 PGTYFCSCPPIYLDDNRSQCDDINECEHNHTCNLQOTCYNLOGGFRCIDPISCCEPYL 300
QY 326 RISDNRCMCPAENPGCRDQPTLLYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 385
DB 301 LISDNRCMCPAENPGCRDQPTLLYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 360
QY 386 GNEGREFYRQRTGPISATLVMTRPIKGRRIOLDLEMITVNTVINFSSVIRLRIYYSQ 445
DB 361 GNEGREFYRQRTGPISATLVMTRPIKGRRIOLDLEMITVNTVINFSSVIRLRIYYSQ 420
QY 446 YPF 448
DB 421 YPF 423

Search completed: July 3, 2003, 17:54:48
Job time : 34.4041 secs

```

QY 65 GYLCTPRTNPVYRGFPYSNPSTPYSGPYPAAPPLSAPNYPRTISRLPICRFYQMDNSQ 124
DB 78 GYLCTPRTNPVYRGFPYSNPSTPYSGPYPAAPPLSAPNYPRTISRLPICRFYQMDNSQ 137
QY 125 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 184
DB 138 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 197
QY 185 PGYSCTCNPGFTLNDGRCSCDVNCAENPCVQCVNTYGSFICRDPGYLEEGVH 244
DB 198 PGYSCTCNPGFTLNDGRCSCDVNCAENPCVQCVNTYGSFICRDPGYLEEGVH 257
QY 245 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQQT 304
DB 258 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQQT 317
QY 305 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPTILYRDMDVSGRSVPAD 364
DB 318 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPTILYRDMDVSGRSVPAD 377
QY 365 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIOLDEMT 424
DB 378 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIOLDEMT 437
QY 425 VNTVINFRGSSVIRLRITYSQYPF 448
DB 438 VNTVINFRGSSVIRLRITYSQYPF 461

RESULT 14
AAY54991
ID AAY54991 standard; Protein; 461 AA.
XX
AC AAY54991;
DT 15-FEB-2000 (first entry)
DE Full length mouse A55 protein sequence.
XX
KW A55 protein; mouse; smooth muscle proliferation; tissue generation;
KW vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;
KW percutaneous transluminal coronary angioplasty; blood coagulation; PRCA;
KW actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
KW tumour metastasis inhibitor; ss.
XX
XX Mus musculus.
XX
XX WO9955864-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1999; 99WO-JP02284.
XX
XX 28-APR-1998; 98JP-0119731.
XX
XX (ONOV ) ONO PHARM CO LTD.
XX
XX Honjo T, Tashiro K, Nakamura T;
XX
XX MPI: 2000-038647/03.
XX
XX N-PSDB; AAZ40031.
XX
XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
XX myoma -
XX
XX Example 5; Page 64-68; 87p; Japanese.
XX
XX This sequence is the mouse A55 protein. The invention relates to the
XX human A55 protein. The protein can be used for the treatment of diseases
XX due to abnormal proliferation of smooth muscle. The polypeptides can be
XX used according their inhibition of the proliferation of vascular smooth

```

```

CC muscle cells, particularly in treating arteriosclerosis or re-narrowing
CC by vascular endothelial thickening after percutaneous transluminal
CC coronary angioplasty (PRCA), or myoma, hematopoietic cell-regulatory
CC activity, cytokine activity, tissue generation/reparation activity,
CC actin/inhibin activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity,
CC cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as
CC a nutrient.
XX
XX Sequence 461 AA:
XX
XX Query Match 93.8%; Score 2376; DB 21; Length 461;
XX Best Local Similarity 93.9%; Pred. No. 1.5e-153;
XX Matches 417; Conservative 10; Mismatches 17; Indels 0; Gaps 0;
QY 5 KRILVTITLALCLPBGNAQAQCTNFGDIDRSGGCLDIDECTIPACRGDMVCVNQNG 64
DB 18 KRILVTITLALMLPHRGNAAQOCTNFGDIDRSGGCLDIDECTIPACRGDMVCVNQNG 77
QY 65 GYLCTPRTNPVYRGFPYSNPSTPYSGPYPAAPPLSAPNYPRTISRLPICRFYQMDNSQ 124
DB 78 GYLCTPRTNPVYRGFPYSNPSTPYSGPYPAAPPLSAPNYPRTISRLPICRFYQMDNSQ 137
QY 125 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 184
DB 138 CVDVDECATDSSHQCNPQTQICINTEGGYTCSTDGWYLLLEGQCLDIDECRYGYCOQLCANV 197
QY 185 PGYSCTCNPGFTLNDGRCSCDVNCAENPCVQCVNTYGSFICRDPGYLEEGVH 244
DB 198 PGYSCTCNPGFTLNDGRCSCDVNCAENPCVQCVNTYGSFICRDPGYLEEGVH 257
QY 245 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQQT 304
DB 258 CSDMDECSFSEFLCOHECVNPGTYFCSCPGYIILLDNRSODINECHRNHTCNLQQT 317
QY 305 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPTILYRDMDVSGRSVPAD 364
DB 318 CYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPCRDQPTILYRDMDVSGRSVPAD 377
QY 365 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIOLDEMT 424
DB 378 IFOMATRRYPGAYYIFQIKSGNEGREFYMRQTPISATLYMTRPIKGPREDIOLDEMT 437
QY 425 VNTVINFRGSSVIRLRITYSQYPF 448
DB 438 VNTVINFRGSSVIRLRITYSQYPF 461

RESULT 15
AAY56751
ID AAY56751 standard; Protein; 423 AA.
XX
XX AAY56751;
XX
XX 22-FEB-2000 (first entry)
XX
XX Smooth muscle proliferation modulating protein mature sequence.
XX
XX Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PRCA;
XX endothelial thickening; percutaneous transluminal coronary angioplasty;
XX myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
XX actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
XX metastasis; nutrient.
XX
XX Mus musculus.
XX
XX WO9955863-A1.
XX
XX 04-NOV-1999.
XX
XX 28-APR-1999; 99WO-JP02283.
XX
XX 28-APR-1998; 98JP-0119731.
XX

```

D7		15-FEB-2000 (first entry)	
XX			
DE		Full length mouse A55 protein sequence.	
XX			
KW	A55 protein; mouse; smooth muscle proliferation; tissue generation;		
KV	vascular smooth muscle cell; arteriosclerosis; tissue reparation; myoma;		
KW	vascular endothelial thickening; haematopoietic cell-regulator; cytokine;		
KW	percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;		
KM	actin; inhiblin; chemotaxis; thrombosis; cadherin; therapy;		
KW	tumour metastasis inhibitor.		
XX			
OS	Mus musculus.		
XX			
PN	MO9955864-A1.		
PD	04-NOV-1999.		
XX			
PF	28-APR-1999; 99WO-JP02284.		
XX			
PR	28-APR-1998; 98JP-0119731.		
XX			
XX	(ONOX) ONO PHARM CO LTD.		
PI	Honjo T, Tashiro K, Nakamura T;		
XX			
DR	WPI: 2000-036647/03.		
DR	N-PDSB: AAZ40029.		
XX			
PT	Novel human polypeptides for treatment of, e.g. arteriosclerosis and		
PT	myoma -		
XX			
PS	Example 4; Page 52-55; 87pp: Japanese.		
XX			
CC	This sequence is the mouse A55 protein. The invention relates to the		
CC	human A55 protein. The protein can be used for the treatment of diseases		
CC	due to abnormal proliferation of smooth muscle. The polypeptides can be		
CC	used according their inhibition of the proliferation of vascular smooth		
CC	muscle cells, particularly in treating arteriosclerosis or re-narrowing		
CC	by vascular endothelial thickening after percutaneous transluminal		
CC	coronary angiolaplasty (PTCA), or myoma, haematopoietic cell-regulatory		
CC	activity, cytokine activity, tissue generation/reparation activity,		
CC	actin/inhibin activity, taxis and chemotaxis activity, blood		
CC	coagulation/thrombotic activity, receptor/ligand activity,		
CC	cadherin/tumour metastasis inhibiting activity; tumour inhibition, and as		
CC	a nutrient.		
XX			
SQ	Sequence 448 AA;		
	Query Match 94.9%; Score 2405; DB 21; Length 448;		
	Best Local Similarity 94.2%; Pred. No. 1.6e-155;		
	Matches 422; Conservative 10; Mismatches 16; Indels 0; Gaps 0		
OY	1 MPGLKRIILVTITLALCLSPGNAQCCTNGFPLDROSGCCLIDECRTIPACRGDMCV 60		
DB	1 MPGLKRIILVTITLALMLPHPGNAQQCTNGFPLDRSGCCLIDECRTIPACRGDMCV 60		
OY	NONGGYLICPTNTNPVRYGRPYSNPYSTPYSGPYPAAPLPASAPNYPITSRPLICRFGYOMD 120		
OY	NONGGYLICPTNTNPVRYGRPYSNPYSTPYSGPYPAAPLPASAPNYPITSRPLICRFGYOMD 120		
DB	61 NONGGYLICPTNTNPVRYGRPYSNPYSTPYSGPYPAAPLPASAPNYPITSRPLICRFGYOMD 120		
OY	ESNOCVDVDECATVDSHOCNPFOICINTBGGYTCSCTDGWLLLEGQCLIDECRYGYCOOL 180		
DB	EGNOCVDVDECATVDSHOCNPFOICINTBGGYTCSCTDGWLLLEGQCLIDECRYGYCOOL 180		
OY	CANVPGSYSCTCNPEFTLNDDRSCODVNECTEMPVCVOTCVNTYSFLICRCDPGELE 240		
DB	CANVPGSYSCTCNPEFTLNDDRSCODVNECTEMPVCVOTCVNTYSFLICRCDPGELE 240		
OY	DGVNCSDMDDESFSSEFLCOHECVNPGYFYFCSPGYILLDNRRSCODINECHRHNHTCN 300		
DB	DGVNCSDMDDESFSSEFLCOHECVNPGYFYFCSPGYILLDNRRSCODINECHRHNHTCT 300		
OY	LQQTTCYNLQAGFKCIDPIRCBEERYLRISDNRCMCAPENFGCRDQPTTILYRDMDVVSGRS 360		

Db	301	SLQTCYNAGGFKCIDPISCEPIYLLIGENRCMAEHTSCDQPFITLIXRMDVYSGRS	360
Qy	361	VPADIFOMQATRRYPGAYYIFQIKSGNGREFRYMOTGPISATLWTRPIKPREIQDL	420
Db	361	VPADIFQMGQATRRYPGAYYIFQIKSGNGREFRYMOTGPISATLWTRPIKPREIQDL	420
Qy	421	EMITVNTYINFRGSSVIRLRIYVSQYF	448
Db	421	EMITVNTYINFRGSSVIRLRIYVSQYF	448
RESULT 13			
ID	AAV56752	standard; Protein: 461 AA.	
XX	AAV56752;		
XX	AAV56752;		
XX	22-FEB-2000	(first entry)	
DE	22-FEB-2000	(first entry)	
XX	Smooth muscle proliferation modulating protein.		
XX	Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;		
KW	endothelial thickening; percutaneous transluminal coronary angioplasty;		
KW	myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;		
KW	actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;		
KW	metastasis; nutrient.		
XX	Mus musculus.		
OS	MO9955863-A1.		
PN	04-NOV-1999.		
XX	04-NOV-1999.		
PD	28-APR-1999;	99MO-JP02283.	
XX	28-APR-1999;	99MO-JP02283.	
PR	28-APR-1998;	98UP-0119731.	
XX	(ONOX) ONO PHARM CO LTD.		
PA	Honjo T, Tashiro K, Nakamura T;		
PI	WPI: 2000-038646/03.		
XX	N-PSDB; AAZ39386; AAZ39387.		
DR	Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma		
XX			
PT	Claim 1; Page 54-56; 70pp; Japanese.		
PS			
XX	The invention provides mouse polypeptides for treatment of diseases due		
CC	to abnormal proliferation of smooth muscle. The polypeptides can be		
CC	produced by standard recombinant methodology. The polypeptides can be		
CC	used according to their inhibition of the proliferation of vascular		
CC	smooth muscle cells, particularly in treating arteriosclerosis or re-		
CC	narrowing by vascular endothelial thickening after percutaneous		
CC	transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-		
CC	regulatory activity, cytokine activity, tissue generation/repair		
CC	activity, actin/inhibitin activity, taxis and chemotaxis activity, blood		
CC	coagulation/thrombotic activity, receptor/ligand activity, cadherin/		
CC	tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.		
CC	The present sequence represents the protein of the invention which can be		
CC	used for modulating smooth muscle cell proliferation.		
SO	Sequence 461 AA;		
Qy	Query Match	93.8%; Score 2376; DB 21; Length 461;	
Db	Best Local Similarity	93.9%; Pred. No. 1.5e-153;	
Db	Matches 417; Conservative 10; Mismatches 17; Indels 0; Gaps 0		
Qy	5 KRITVTITLALCLSPGNAQACTNGFPLDRSGGCLDIDERTTPACRGDMKCVNONG	64	
Db	18 RMILITVITLALMLPBGNAQOCTNGFPLDRSGGCLDIDERTTPACRGDMKCVNONG	77	

CC can be modulated with agents that bind to these polypeptides, resulting
 CC in the death of the cells. The polynucleotides encoding these
 CC polypeptides are useful in the recombinant production of the
 CC polypeptides, as a hybridisation probe to screen libraries to isolate
 CC homologous sequences, or to map the gene. They may also be used for
 CC analysing genetic disorders, and to produce transgenic animals which are
 CC useful for the development and screening of therapeutically useful
 CC reagents. The polynucleotides can also be used in gene therapy e.g. to
 CC replace a defective gene.

SO Sequence 448 AA:

Query Match 99.8%; Score 2527; DB 22; Length 448;
 Best Local Similarity 99.8%; Pred. No. 8.1e-164;
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MPGIKRITVTITLALCLPSGNAOACTNGFDLRSGCGLDIDECRTIPEACRGDMCV 60
Db 1 MPGIKRITVTITLALCLPSGNAOACTNGFDLRSGCGLDIDECRTIPEACRGDMCV 60
OY 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
Db 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
OY 121 ESNOCVVDDECATSHOCNPTQICINTEGGYCTSDGWLLEGGCLDIDECRYGCOOL 180
Db 121 ESNOCVVDDECATSHOCNPTQICINTEGGYCTSDGWLLEGGCLDIDECRYGCOOL 180
OY 181 CANPGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
Db 181 CANPGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
OY 241 DGVCSDMDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSQDINECHRNHTCN 300
Db 241 DGVCSDMDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSQDINECHRNHTCN 300
OY 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
Db 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
OY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQTGPISATLVMTPIKGPRIQDL 420
Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQTGPISATLVMTPIKGPRIQDL 420
OY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

```

RESULT 11

AAV56750 ID AAV56750 standard; Protein: 448 AA.

AAV56750:

22-FEB-2000 (first entry)

Smooth muscle proliferation modulating protein.

CC Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA;
 CC endothelial thickening; percutaneous transluminal coronary angioplasty;
 CC myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
 CC actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
 CC metastasis; nutrient.

XX Mus musculus.

XX W09555863-A1.

XX 04-NOV-1999.

XX 28-APR-1999; 99WO-JP02283.

XX

PR 28-APR-1998; 98JP-0119731.

XX (ONOX) ONO PHARM CO LTD.

XX Honjo T, Tashiro K, Nakamura T;

XX WPI: 2000-038646/03.

XX N-PSDB: AA239383, AA239384.

XX Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

XX Claim 1; Page 42-44; 70pp; Japanese.

CC The invention provides mouse polypeptides for treatment of diseases due
 CC to abnormal proliferation of smooth muscle. The polypeptides can be
 CC produced by standard recombinant methodology. The polypeptides can be
 CC used according to their inhibition of the proliferation of vascular
 CC smooth muscle cells, particularly in treating arteriosclerosis or re-
 CC narrowing by vascular endothelial thickening after percutaneous
 CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
 CC regulatory activity, cytokine activity, tissue generation/repair
 CC activity, actin/inhibitor activity, taxis and chemotaxis activity, blood
 CC coagulation/chromoblastic activity, receptor/ligand activity, cadherin/
 CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
 CC The present sequence represents the protein of the invention which can be
 CC used for modulating smooth muscle cell proliferation.

SO Sequence 448 AA:

Query Match 94.9%; Score 2405; DB 21; Length 448;
 Best Local Similarity 94.2%; Pred. No. 1.6e-155;
 Matches 422; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

```

OY 1 MPGIKRITVTITLALCLPSGNAOACTNGFDLRSGCGLDIDECRTIPEACRGDMCV 60
Db 1 MPGIKRITVTITLALCLPSGNAOACTNGFDLRSGCGLDIDECRTIPEACRGDMCV 60
OY 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
Db 61 NONGYLCIPRTNPNVYRGPSNPNSTPYSGPYPAAPPLSAPNPTISRLICRFGYQMD 120
OY 121 ESNOCVVDDECATSHOCNPTQICINTEGGYCTSDGWLLEGGCLDIDECRYGCOOL 180
Db 121 ESNOCVVDDECATSHOCNPTQICINTEGGYCTSDGWLLEGGCLDIDECRYGCOOL 180
OY 181 CANPGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
Db 181 CANPGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
OY 241 DGVCSDMDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSQDINECHRNHTCN 300
Db 241 DGVCSDMDECSFSEFLCOHECVNPGTYFCSCPPGYILLDNRSQDINECHRNHTCN 300
OY 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
Db 301 LQOTCYNLQGGFKCIDPIRCEPEYLRISDNRCMPAENPCGRDQPTIILYRDMVYSGRS 360
OY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQTGPISATLVMTPIKGPRIQDL 420
Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQTGPISATLVMTPIKGPRIQDL 420
OY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

```

RESULT 12

AAV54990 ID AAV54990 standard; Protein: 448 AA.

AAV54990:

XX

CC the proliferation or differentiation of chondrocyte cells. The PRO
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 448 AA:

Query Match 99.8%; Score 2527; DB 22; Length 448;
 Best Local Similarity 99.8%; Pred. No. 8,1e-164;
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGIKRLITVTIALCLPSPGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
 DB 1 MGIKRLITVTIALCLPSPGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
 QY 61 NONGGYLCIRTNVYRGPSNPTSPYSGPYPAAPPLSAPNPTISRPILCRFGYQMD 120
 DB 61 NONGGYLCIRTNVYRGPSNPTSPYSGPYPAAPPLSAPNPTISRPILCRFGYQMD 120
 QY 121 ESNQCVVDECATDSDHQCNPQICINTEGYTSCSTDGWMLEGGCLDIDECRYGCOOL 180
 DB 121 ESNQCVVDECATDSDHQCNPQICINTEGYTSCSTDGWMLEGGCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRDPGELEE 240
 DB 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRDPGELEE 240
 QY 241 DGVHCSMDCESEFELCQHCVCVNOPTGYFSCSPRGYTLILDNRSCODINEEHRHNTN 300
 DB 241 DGVHCSMDCESEFELCQHCVCVNOPTGYFSCSPRGYTLILDNRSCODINEEHRHNTN 300
 QY 301 LQOTCYNIAGGFKCIDPRCEPFLRISDNRCMCAPENPCGDOFTILYRMDVYSGRS 360
 DB 301 LQOTCYNIAGGFKCIDPRCEPFLRISDNRCMCAPENPCGDOFTILYRMDVYSGRS 360
 QY 361 VPADIFQMAATRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKPREIQLDL 420
 DB 361 VPADIFQMAATRYPGAYYIFQIKSGNEGREFYMRQTPISATLVMTRPKPREIQLDL 420
 QY 421 EMITVNTVINRGSSVIRLRIVYQYPR 448
 DB 421 EMITVNTVINRGSSVIRLRIVYQYPR 448

RESULT 10

ID AAB31183 standard; Protein: 448 AA.

XX AAB31183:

DT 20-APR-2001 (first entry)

XX Amino acid sequence of human polypeptide PRO210.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
 KW PRO185; PRO210; PRO215; PRO242; PRO288; PRO365; PRO361; PRO1308;
 KW PRO1183; PRO1272; PRO1419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
 KW PRO1600; PRO9940; PRO553; PRO301; PRO187; PRO337; PRO4111; PRO4356;
 KW PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO2630;
 KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..25 /note- "signal peptide"
 FT Modified-site 21..27 /note- "N-myristoylation site"
 FT Binding-site 54..57 /note- "cell attachment site"
 FT Modified-site 64..70

FT /note- "N-myristoylation site"
 FT 144..156
 FT /note- "aspartic acid and asparagine hydroxylation site"
 FT 149..155
 FT /note- "N-myristoylation site"
 FT 186..192
 FT /note- "N-myristoylation site"
 FT 226..232
 FT /note- "N-myristoylation site"
 FT 242..248
 FT /note- "N-myristoylation site"
 FT 267..273
 FT /note- "N-myristoylation site"
 FT 283..287
 FT /note- "N-glycosylation site"
 FT 296..300
 FT /note- "N-glycosylation site"
 FT 310..316
 FT /note- "N-myristoylation site"

WO20007037-A2.

PD 21-DEC-2000.

PE 22-MAY-2000; 2000WO-US14042.

XX 15-JUN-1999; 99US-0139695.
 PR 20-JUL-1999; 99US-0145070.
 PR 26-JUL-1999; 99US-0145698.
 PR 17-AUG-1999; 99US-0149396.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28565.
 PR 07-DEC-1999; 99US-0169495.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 15-MAY-2000; 2000WO-US13358.
 PR 17-MAY-2000; 2000WO-US13705.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
 PI Ferrara N, Fong S, Gao W, Gerber H, Gerltzen ME, Goddard A;
 PI Godowski PJ, Gunney AL, Kljavin IJ, Mather JP, Napier MA, Pan J;
 PI Paoni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;
 PI Wood WI, Zhang Z;

XX WPI: 2001-050091/06.
 DR N-PSDB: AAC86968.

PT Isolated nucleic acid molecule encoding a PRO polypeptide which is a
 PT transmembrane polypeptide is useful for gene therapy and identification
 PT of related polypeptides -

PS Claim 12; Fig 10; 244p; English.

XX The present sequence represents a human secreted and transmembrane
 CC polypeptide. The specification describes human polypeptides, designated
 CC PRO196, PRO444, PRO183, PRO185, PRO210, PRO215, PRO242, PRO288,
 CC PRO365, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
 CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO553, PRO301, PRO187,
 CC PRO337, PRO4111, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
 CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells

Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic thyroiditis), diabetes mellitus, immune-mediated renal disease (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic inflammatory demyelinating polyneuropathy, infectious hepatitis (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune hepatitis, primary biliary cirrhosis, granulomatous hepatitis, and sclerosing cholangitis), gluten-sensitive enteropathy, and ulcerative colitis (Crohn's disease), gluten-mediated skin diseases including Whipple's disease, autoimmune or immune-mediated dermatitis, psoriasis, bullous skin diseases, erythema multiforme, contact dermatitis, and asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity, urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity pneumonitis, and transplantation associated diseases (graft rejection, and graft-versus-host-disease). (1), its antagonists or fragment can also be used as an adjuvant in treatment of tumors.

Antibodies against (1) can also be used for diagnosing such diseases.

This sequence represents a human EGF-like homologue encoded by cDNA clone DNA32279 which is described in the invention.

Sequence 448 AA:

Query Match 99.88; Score 2527; DB 20; Length 448;
Best Local Similarity 99.88; Pred. No 8.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MPEIKRLVTITIALCLPSGNAOCTNGFDLDROSGQLDIDECRTIPEACRQDMCV 60
1 MPEIKRLVTITIALCLPSGNAOCTNGFDLDROSGQLDIDECRTIPEACRQDMCV 60
61 NONGGYLCIRTNVYRGPSNPSTPSGYPAPAAPLSAPNYTISRPLCRGYOND 120
61 NONGGYLCIRTNVYRGPSNPSTPSGYPAPAAPLSAPNYTISRPLCRGYOND 120
61 NONGGYLCIRTNVYRGPSNPSTPSGYPAPAAPLSAPNYTISRPLCRGYOND 120
121 ESNOCVDECATSHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
121 ESNOCVDECATSHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
121 ESNOCVDECATSHOCNPTQICINTEGTYCTSDGYWLEGGCLDIDECRYGCOOL 180
181 CANVPGSYCTCNGPFTLNEDGRSCQDVNECAPCVQTCVNTGSGFCRCDPGYLEE 240
181 CANVPGSYCTCNGPFTLNEDGRSCQDVNECAPCVQTCVNTGSGFCRCDPGYLEE 240
181 CANVPGSYCTCNGPFTLNEDGRSCQDVNECAPCVQTCVNTGSGFCRCDPGYLEE 240
241 DGVHCSMDSECFSEFLCQHECVNPGTYFCSCPGYITLLDNRSCQDINECHRNHTCN 300
241 DGVHCSMDSECFSEFLCQHECVNPGTYFCSCPGYITLLDNRSCQDINECHRNHTCN 300
241 DGVHCSMDSECFSEFLCQHECVNPGTYFCSCPGYITLLDNRSCQDINECHRNHTCN 300
301 LQOTCYNLOGGFKCIDPICEPEYLRISDNRCMCPAENGCRCRDPFTILYRDMDVYSGRS 360
301 LQOTCYNLOGGFKCIDPICEPEYLRISDNRCMCPAENGCRCRDPFTILYRDMDVYSGRS 360
301 LQOTCYNLOGGFKCIDPICEPEYLRISDNRCMCPAENGCRCRDPFTILYRDMDVYSGRS 360
361 VPADIFQWQATRRPGATYITQIKSGNGREFYRGTGPIATVTRPIKGRREIOLDL 420
361 VPADIFQWQATRRPGATYITQIKSGNGREFYRGTGPIATVTRPIKGRREIOLDL 420
361 VPADIFQWQATRRPGATYITQIKSGNGREFYRGTGPIATVTRPIKGRREIOLDL 420
421 EMITVNTYINFRGSSVIRLRIVSQYPF 448
421 EMITVNTYINFRGSSVIRLRIVSQYPF 448
421 EMITVNTYINFRGSSVIRLRIVSQYPF 448

RESULT 9

AAU29227 standard; Protein: 448 AA.

AAU29227;

18-DEC-2001 (first entry)

Human PRO polypeptide sequence #204.

PRO polypeptide: mammal; tumour; cancer; human; cattle; horse; sheep; dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX Homo sapiens.

XX WO200168848-A2.

XX 20-SEP-2001.

XX 28-FEB-2001; 2001MO-US06520.

01-MAR-2000; 2000MO-US05601.
02-MAR-2000; 2000MO-US05841.
03-MAR-2000; 2000US-187202P.
06-MAR-2000; 2000US-186968P.
14-MAR-2000; 2000US-189320P.
14-MAR-2000; 2000US-189328P.
15-MAR-2000; 2000MO-US06884.
21-MAR-2000; 2000US-190828P.
21-MAR-2000; 2000US-191007P.
21-MAR-2000; 2000US-191048P.
21-MAR-2000; 2000US-191314P.
28-MAR-2000; 2000US-192655P.
29-MAR-2000; 2000US-193032P.
30-MAR-2000; 2000US-193053P.
04-APR-2000; 2000MO-US08439.
04-APR-2000; 2000US-194449P.
11-APR-2000; 2000US-195975P.
11-APR-2000; 2000US-196008P.
11-APR-2000; 2000US-196187P.
11-APR-2000; 2000US-196690P.
11-APR-2000; 2000US-196820P.
18-APR-2000; 2000US-198121P.
18-APR-2000; 2000US-198585P.
25-APR-2000; 2000US-199397P.
25-APR-2000; 2000US-199550P.
25-APR-2000; 2000US-199634P.
03-MAY-2000; 2000US-201516P.
17-MAY-2000; 2000MO-US13705.
22-MAY-2000; 2000MO-US14042.
30-MAY-2000; 2000MO-US14941.
02-JUN-2000; 2000MO-US15264.
05-JUN-2000; 2000US-209832P.
28-JUL-2000; 2000MO-US20710.
28-AUG-2000; 2000US-054848P.
24-AUG-2000; 2000MO-US23328.
08-NOV-2000; 2000MO-US30952.
01-DEC-2000; 2000MO-US32678.
20-DEC-2000; 2000MO-US34956.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL, Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI: 2001-602746/68.
N-PSDB: AAS46128.

Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 11; Fig 408; 774pp: English.

Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate

PN US2001051358-A1.
 XX 13-DEC-2001.
 XX 25-MAR-1999; 99US-0275805.
 XX 11-APR-1997; 97US-0839525.
 XX 10-APR-1996; 96MO-US05247.
 PA (OLSEN/) OLSEN H S.
 PA (LIHH/) LI H.
 PI Olsen HS, LI H;
 DR N-PSDB; ABR13627.
 DR WPI; 2002-121417/16.
 PT New nucleic acid encoding human extracellular/epidermal growth factor,
 PT useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
 PT also related polypeptides -
 XX Claim 10; Fig 1; 22pp; English.
 CC The invention relates to a novel polynucleotide which is at
 CC least 95% identical with a sequence (ATCC 97285) encoding mature human
 CC extracellular protein-like/epidermal growth factor (EGF)-like protein.
 CC EGF. Also included are the EGF EGF domains, a vector containing
 CC the polynucleotide, a host cell containing the vector, anti-EGF
 CC antibodies and antagonists of EGF. The polynucleotide is used for
 CC recombinant production of EGF, in gene therapy, as hybridisation probes,
 CC as antisense antagonists and for chromosome identification. The protein
 CC is used to treat patients who require EGF, to identify specific
 CC antagonists, used to treat conditions that require inhibition of EGF
 CC (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound
 CC healing, neurological trauma, acquired immunodeficiency syndrome
 CC (AIDS)-related dementia, ocular disorders, kidney disorders, liver
 CC disorders, hair follicle growth promotion, burns, ulcers, corneal
 CC incisions, corneal inflammation, neoplasms and psoriasis), to raise
 CC specific antibodies and to characterise receptors. The present
 CC sequence represents EGF.
 CC
 SO Sequence 448 AA;
 Query Match 100.0%; Score 2533; DB 23; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MPGKRLITVITIALCLPSPGNAOCTNGFDLDRSGCCLDIDERTIPEACRGDMCV 60
 1 MPGKRLITVITIALCLPSPGNAOCTNGFDLDRSGCCLDIDERTIPEACRGDMCV 60
 61 NONGVYLCIPRTNPNVRYRGPYNSPYSGPYPAAPPLSAPNPTISPLICRFYQMD 120
 61 NONGVYLCIPRTNPNVRYRGPYNSPYSGPYPAAPPLSAPNPTISPLICRFYQMD 120
 61 NONGVYLCIPRTNPNVRYRGPYNSPYSGPYPAAPPLSAPNPTISPLICRFYQMD 120
 121 ESNQCVVDDECATDSDHQCNPQICINTEGGYTCSTDGYMLLEGQCLDIDECRYGVCQL 180
 121 ESNQCVVDDECATDSDHQCNPQICINTEGGYTCSTDGYMLLEGQCLDIDECRYGVCQL 180
 121 ESNQCVVDDECATDSDHQCNPQICINTEGGYTCSTDGYMLLEGQCLDIDECRYGVCQL 180
 181 CANVPGSYSCNCGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 181 CANVPGSYSCNCGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 181 CANVPGSYSCNCGFTLNEDGRSCQDVNECATENPCVOTCVNTYSFICRCDPGYLEE 240
 241 DGVMCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRHNHCN 300
 241 DGVMCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRHNHCN 300
 241 DGVMCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSCQDINECHRHNHCN 300
 301 LQGTQYNLQGGFKCIDPICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 301 LQGTQYNLQGGFKCIDPICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 301 LQGTQYNLQGGFKCIDPICEEPYLRISDNRCMCPAENGCARDQPTILYRDMVYSGRS 360
 361 VPADIFQWQATRRPGAYITRQIKSGNREGRYMRQGTGISTATVTRPIKGRREIQDL 420
 361 VPADIFQWQATRRPGAYITRQIKSGNREGRYMRQGTGISTATVTRPIKGRREIQDL 420

DB 361 VPADIFQWQATRRPGAYITRQIKSGNREGRYMRQGTGISTATVTRPIKGRREIQDL 420
 QY 421 EMITVNTVINRGSSVIRLRITVYSQYF 448
 DB 421 EMITVNTVINRGSSVIRLRITVYSQYF 448
 RESULT 8
 AAY08063
 ID AAY08063 standard; Protein: 448 AA.
 AC AAY08063;
 XX 11-SEP-2000 (first entry)
 DT Human EGF-like homologue protein (PRO217) encoded by DNA32279 cDNA.
 DE
 XX Inflammatory cell infiltration; immune response; T cell proliferation;
 KW anti-inflammatory; anti-autoimmune; anti-diabetic; spondyloarthritis;
 KW T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
 KW inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
 KW diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
 KW multiple sclerosis; polynuropathy; hepatitis; cirrhosis; enteropathy;
 KW sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
 KW skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
 KW food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
 KW idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
 EGF-like.
 OS Homo sapiens.
 XX W05914241-A2.
 XX 25-MAR-1999;
 XX 17-SEP-1998; 98MO-US19437.
 XX 17-SEP-1997; 97US-0059119.
 XX 18-SEP-1997; 97US-0059263.
 XX 28-OCT-1997; 97US-0063550.
 XX 12-NOV-1997; 97US-0065186.
 XX 21-NOV-1997; 97US-0066364.
 XX 24-NOV-1997; 97US-0066770.
 XX 04-JUN-1998; 98US-0088026.
 PA (GETH) GENENTECH INC.
 PI Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;
 DR WPI; 1999-229499/19.
 DR N-PSDB; AAX37670.
 PT Composition containing novel polypeptide PRO245, its agonist or
 PT antagonist -
 XX Example 1; Fig 6a; 177pp; English.
 PS This invention describes a novel composition containing (apart from a
 CC carrier or excipient), a novel PRO245 polypeptide (I), its agonist or
 CC antagonist, or their fragments, for modulating: (i) infiltration of
 CC inflammatory cells into tissue; (ii) an immune response; or (iii) T cell
 CC proliferation. The products of the invention have anti-inflammatory,
 CC (1)-(iii). The products of the invention have anti-inflammatory,
 CC anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists
 CC and their fragments, are used to treat immune-related diseases,
 CC particularly T cell-mediated diseases. The diseases treated include
 CC systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic
 CC arthritis, spondyloarthritis, systemic sclerosis (scleroderma),
 CC idiopathic inflammatory myopathies (dermatomyositis, polymyositis),
 CC Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune
 CC hemolytic anemia (immune pancytopenia, paroxysmal nocturnal
 CC hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic
 CC purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,

Db 241 DGVHCSMDCECSSEFLCOHECVNOPTGYCSCPBYILLDDNRS CODINECEHRNHCN 300
QY 301 LOOTCYNLOGGFRCIDPICEEPEYLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFRCIDPICEEPEYLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQGTGPISATLVMTRPIKGPRIQDL 420
Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQGTGPISATLVMTRPIKGPRIQDL 420
QY 421 EMITVNTVINFGRSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINFGRSSVIRLRIVYSQYPF 448

RESULT 6
ID AAM93573 standard; Protein: 448 AA.
AC AAM93573;
XX 06-NOV-2001 (first entry)
DE Human polypeptide, SEQ ID NO: 3357.
XX Human polypeptide, SEQ ID NO: 3357.
KM Human, full length cDNA; cDNA synthesis; oligo-capping.
XX Homo sapiens.
OS Homo sapiens.
XX EPI130094-A2.
PN EPI130094-A2.
PD 05-SEP-2001.
XX 07-JUL-2000; 2000EP-0114089.
PE 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX (HELI-) HELIX RES INSTR.
PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX MPI: 2001-524255/58.
DR N-PSDB: AAK94505.
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX Claim 8: SEQ ID NO 3357; 1380bp + sequence listing; English.
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 448 AA:
Query Match 100.0%; Score 2533; DB 22; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.2e-164;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MFGIKRLVTITLALCLPSPGNAOACTNGFDLDROSGQCIDIDECRTIPEACRGDMNCV 60
|||||

Db 1 MFGIKRLVTITLALCLPSPGNAOACTNGFDLDROSGQCIDIDECRTIPEACRGDMNCV 60
QY 61 NONGGYLCLIPRTNPPYRGPNYSNPYSRPGYPAAAPPLSAPNPPTISRPLICRGYOMD 120
Db 61 NONGGYLCLIPRTNPPYRGPNYSNPYSRPGYPAAAPPLSAPNPPTISRPLICRGYOMD 120
QY 121 ESNOCYDVDECATDSHOCNPTQICINTEGGYTCSTDGWYLLBEGCCLDIDECRGYCOOL 180
Db 121 ESNOCYDVDECATDSHOCNPTQICINTEGGYTCSTDGWYLLBEGCCLDIDECRGYCOOL 180
QY 181 CANPESYSTCTCNPGFTLNDGNSQDYNBCATNPCVQTCVNTYNGSTICRCDCPELEB 240
Db 181 CANPESYSTCTCNPGFTLNDGNSQDYNBCATNPCVQTCVNTYNGSTICRCDCPELEB 240
QY 241 DGVHCSMDCECSSEFLCOHECVNOPTGYCSCPBYILLDDNRS CODINECEHRNHCN 300
Db 241 DGVHCSMDCECSSEFLCOHECVNOPTGYCSCPBYILLDDNRS CODINECEHRNHCN 300
QY 301 LOOTCYNLOGGFRCIDPICEEPEYLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFRCIDPICEEPEYLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
QY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQGTGPISATLVMTRPIKGPRIQDL 420
Db 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFFYMRQGTGPISATLVMTRPIKGPRIQDL 420
QY 421 EMITVNTVINFGRSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINFGRSSVIRLRIVYSQYPF 448

RESULT 7
ID AAM75494 standard; Protein: 448 AA.
AC AAM75494;
XX 23-APR-2002 (first entry)
DE Human extracellular protein-like/EGF-like protein, EGF.
XX Human; extracellular protein-like protein; EGF-like;
KM protein; epidermal growth factor; EGF; ARCC 97285; gene therapy;
KM vascular smooth muscle cell proliferation; Marfan syndrome;
KM wound healing; neurological trauma; acquired immunodeficiency syndrome;
KM AIDS-related dementia; ocular disorder; kidney disorder; liver disorder;
KM hair follicle growth promotion; burn; ulcer; corneal incision;
KM corneal inflammation; neoplasm; psoriasis.
XX Homo sapiens.
OS Homo sapiens.
XX
XX Key
FH 1..25 Location/Qualifiers
FT /label= signal_peptide
FT 26..448
FT /label= Mature_EGF
FT /note= "This region is specifically claimed in claim 10"
FT 112..153
FT /label= EGF_1_domain
FT /note= "This domain is specifically claimed in claim 10"
FT 154..190
FT /label= EGF_2_domain
FT /note= "This domain is specifically claimed in claim 10"
FT 191..230
FT /label= EGF_3_domain
FT /note= "This domain is specifically claimed in claim 10"
FT 231..271
FT /label= EGF_4_domain
FT /note= "This domain is specifically claimed in claim 10"
FT 272..314
FT /label= EGF_5_domain
FT /note= "This domain is specifically claimed in claim 10"
XX

PT smooth muscle cell growth, vasculogenesis, restenosis or
 PT atherosclerosis
 XX
 PS Claim 4: Page 46-47; 49pp; English.

CC This is the amino acid sequence of the human secreted protein AK647. The
 CC polynucleotide sequence was obtained from a human foetal kidney cDNA
 CC library. AK647 homologues in chicks and rodents are involved in aortic
 CC tissue development. The spatial and temporal distribution of AK647
 CC indicated that it acts as a modulator of smooth muscle cells in
 CC vasculogenesis during embryonic development. The primary structure of
 CC AK647 consists of multiple EGF domains. The AK647 protein can be used as
 CC a nutritional source or supplement. The protein shows both inhibitory and
 CC inducing, cytokine, cell proliferation and cell differentiation activity.
 CC The protein may also be used in the treatment of immune deficiencies and
 CC disorders, including severe combined immunodeficiency (SCID), HIV and
 CC other viral, bacterial and fungal infections. Regulation of immune
 CC responses may also be carried out by the AK647 protein. Other uses of the
 CC protein include a role in the regulation of haematopoiesis and in the
 CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
 CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
 CC as well as for wound healing and in the treatment of ulcers and burns.
 CC The polynucleotides and proteins can be used for preventing, treating or
 CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
 CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
 CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
 CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activity. AK647 specific
 CC antibodies can be used for promoting smooth muscle cell growth or
 CC vasculogenesis. The proteins and polynucleotides can also be used for
 CC detection, diagnosis and drug screening.

XX Sequence 448 AA;

Query Match 100.0%; Score 2533; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 3.2e-164; Indels 0; Gaps 0;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGKIRLITVTITLALCLPSGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 DB 1 MPGKIRLITVTITLALCLPSGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 QY 61 NONGGYLCIRPTNRYRGPYSNPYSTPGYPAAPPLSAPNYPITSRPLICRFGYQMD 120
 DB 61 NONGGYLCIRPTNRYRGPYSNPYSTPGYPAAPPLSAPNYPITSRPLICRFGYQMD 120
 QY 121 ESNOCVDVDECATDSHOCPNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 DB 121 ESNOCVDVDECATDSHOCPNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 DB 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 QY 301 LOOTCYMLGGGFKCIDIPRCEEPYLRISDNRCMPAENPGCRDOPFTIILYRDMVYSGNS 360
 DB 301 LOOTCYMLGGGFKCIDIPRCEEPYLRISDNRCMPAENPGCRDOPFTIILYRDMVYSGNS 360
 QY 361 VPADIFQOMATTRYPGAYYIFQIRSGNREFYMRGTGPISATLVTRPIKGPRIEQLDL 420
 DB 361 VPADIFQOMATTRYPGAYYIFQIRSGNREFYMRGTGPISATLVTRPIKGPRIEQLDL 420
 QY 421 EMITVNTVINFRGSSVIRLITYSOYR 448
 DB 421 EMITVNTVINFRGSSVIRLITYSOYR 448

RESULT 5
 AAY54989

ID AAY54989 standard; Protein; 448 AA.

XX AAY54989;

XX 15-FEB-2000 (first entry)

XX Full length human A55 protein sequence.

DE A55 protein; human; smooth muscle proliferation; tissue generation;
 KW vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
 KW vascular endothelial thickening; haematopoietic cell-regulator; cytokine;
 KW percutaneous transluminal coronary angioplasty; blood coagulation; PRCA;
 KW actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
 KW tumour metastasis inhibitor.

OS Homo sapiens.

PN W09955864-A1.

XX 04-NOV-1999.

XX 28-APR-1999; 99WO-JP02284.

XX 28-APR-1998; 98JP-0119731.

XX (ONOX) ONO PHARM CO LTD.

PI Honjo T, Tashiro K, Nakamura T;

DR WPI: 2000-038647/03.

DR N-PSDB: AAZ40027.

XX

XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
 PT myoma -

PS Claim 1: Page 76-80; 87pp; Japanese.

CC This sequence is the human A55 protein of the invention. The protein
 CC can be used for the treatment of diseases due to abnormal proliferation
 CC of smooth muscle. The polypeptides can be used according their inhibition
 CC of the proliferation of vascular smooth muscle cells, particularly in
 CC treating arteriosclerosis or re-narrowing by vascular endothelial
 CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
 CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
 CC tissue generation/repairation activity, actin/inhibin activity, taxis
 CC and chemotaxis activity, blood coagulation/thrombotic activity,
 CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
 CC tumour inhibition, and as a nutrient.

XX Sequence 448 AA;

Query Match 100.0%; Score 2533; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 3.2e-164; Indels 0; Gaps 0;

Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGKIRLITVTITLALCLPSGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 DB 1 MPGKIRLITVTITLALCLPSGNAQAQCTNGFDLDRSGGCLDIDERTIPEACRGDMCV 60
 QY 61 NONGGYLCIRPTNRYRGPYSNPYSTPGYPAAPPLSAPNYPITSRPLICRFGYQMD 120
 DB 61 NONGGYLCIRPTNRYRGPYSNPYSTPGYPAAPPLSAPNYPITSRPLICRFGYQMD 120
 QY 121 ESNOCVDVDECATDSHOCPNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 DB 121 ESNOCVDVDECATDSHOCPNPTQICINTEGGYTCSTDGWMLLEGQCLDIDECRYGCOOL 180
 QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYELE 240
 QY 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300
 DB 241 DGVHCSMDDECSFSEFLCQHECVNPGTYFCSPGTYILLDNRSQDINECEHRNHTCN 300

Oy	61	NONGCVLICIPRTNPVYRGYSNPSPYSGPRPAAPPLSAPRYPTISRLICRFGYQMD	120
Dd	61	NONGGYLCIPRTNPVYRGYSNPSPYSGPRPAAPPLSAPRYPTISRLICRFGYQMD	120
Oy	121	ESNQCVDVDECAFTDShOCNPQTICINTEGGYTCSCIDGYWMLLEGQCLDIDECRYGCQQL	180
Dd	121	ESNQCVDVDECAFTDShOCNPQTICINTEGGYTCSCIDGYWMLLEGQCLDIDECRYGCQQL	180
Oy	181	CANPFGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRCDPGYLEE	240
Dd	181	CANPFGYSCTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRCDPGYLEE	240
Oy	241	DGVHCSMDDECSFSEFLCOHECVNOPGTYFGSCGPPEYIILDDNRSCODINECHRNHTCN	300
Dd	241	DGVHCSMDDECSFSEFLCOHECVNOPGTYFGSCGPPEYIILDDNRSCODINECHRNHTCN	300
Oy	301	LQOTCYNLQGGFKCIDPIRCEPEYLRISDRNCMPAENPGCRDOPFTIILYRDMDVYSGRS	360
Dd	301	LQOTCYNLQGGFKCIDPIRCEPEYLRISDRNCMPAENPGCRDOPFTIILYRDMDVYSGRS	360
Oy	361	VPADIFQOATTRYRGAYITFOIKSGNREERYMROTGFIATLVNTRPIKGFREIQLDL	420
Dd	361	VPADIFQOATTRYRGAYITFOIKSGNREERYMROTGFIATLVNTRPIKGFREIQLDL	420
Oy	421	EMITVNTVINFRGSSVIRLRIYVSQYF	448
Dd	421	EMITVNTVINFRGSSVIRLRIYVSQYF	448
RESULT 3			
ID	AAW94281	standard; Protein; 448 AA.	
AC	AAW94281;		
DT	07-MAY-1999	(first entry)	
DE	Human extracellular matrix protein (ECMP)-1.		
KW	Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;		
XX	Immunologic disorder; human.		
XX	Homo sapiens.		
XX	OS		
XX	PN	MO9900410-A2.	
XX	DD	07-JAN-1999.	
XX	PF	23-JUN-1998; 98WO-US13012.	
XX	PR	27-JUN-1997; 97US-0884072.	
XX	PA	(INCY-) INCYTE PHARM INC.	
XX	PI	Bandman O, Corley NC, Guegler KJ;	
XX	DR	WPI: 1999-095674/08.	
XX	PT	N-PDB: AAX05359.	
XX	PT	New polynucleotide encoding extracellular matrix protein, ECMP-1 -	
XX	PT	disorders and cancer	
XX	PS	Claim 1; Fig 1A-G; 79pp; English.	
CC	CC	This represents a human extracellular matrix protein (ECMP)-1. Host	
CC	CC	cells containing a vector comprising the ECMP-1 nucleic acid are used	
CC	CC	for the recombinant production of the protein. ECMP-1 and its	
CC	CC	(ant)agonists, are useful in the diagnosis, prevention, and treatment	
CC	CC	of cancer and immune disorders.	
XX	XX	Sequence 448 AA:	

Query Match	Similarity	100.0%	Score 2533	DB 20	Length 448
Best Local	Similarity	100.0%	Pred. No. 3.2e-164		
Matches 448	Conservative	0	Mismatches 0	Indels 0	Gaps 0
QY	1	MPGIRLITVITLALCLPSPGNAQACNTNGFPLDRQSGCLDIDECRTIPEACRGDMVCY	60		
DB	1	MPGIRLITVITLALCLPSPGNAQACNTNGFPLDRQSGCLDIDECRTIPEACRGDMVCY	60		
QY	61	NONGYLICIPRTNPYRGPYSNPSTPYSGPYAAPAPLSABNPPTISRLICRFGYQMD	120		
DB	61	NONGYLICIPRTNPYRGPYSNPSTPYSGPYAAPAPLSABNPPTISRLICRFGYQMD	120		
QY	121	ESNOCVVDDECATSHQCNPTQICINTEGGTYCCTDGTWLLGSCCLDIDECRYGCOOL	180		
DB	121	ESNOCVVDDECATSHQCNPTQICINTEGGTYCCTDGTWLLGSCCLDIDECRYGCOOL	180		
QY	181	CANVPGSYSCTCNPGFTLNEDGRSCQDYNCEATEMPCVQTCVNTGSGFICRCDPGYELEB	240		
DB	181	CANVPGSYSCTCNPGFTLNEDGRSCQDYNCEATEMPCVQTCVNTGSGFICRCDPGYELEB	240		
QY	241	DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPBGYITLLDNRSCODINECHRNHTCN	300		
DB	241	DGVHCSMDDECSFSEFLQHECVNPGTYFCSCPBGYITLLDNRSCODINECHRNHTCN	300		
QY	301	LQQTCTNLTQGGFKCIDPRICEBPYLRISDNRCMCAENPGCNDQFTLLYRMDVVSGRS	360		
DB	301	LQQTCTNLTQGGFKCIDPRICEBPYLRISDNRCMCAENPGCNDQFTLLYRMDVVSGRS	360		
QY	361	VPADIFQWQATTRYPGAYYIFQIKSGNEGREFYMQTGPISATLVMTFRPKGPRETQDL	420		
DB	361	VPADIFQWQATTRYPGAYYIFQIKSGNEGREFYMQTGPISATLVMTFRPKGPRETQDL	420		
QY	421	EMITVNTVINERGSVIRLRITVYSQYPF 448			
DB	421	EMITVNTVINERGSVIRLRITVYSQYPF 448			
RESULT 4					
ID	AAV57058	AAV57058 standard; Protein; 448 AA.			
AC	AAV57058;				
DT	21-FEB-2000	(first entry)			
DE	Amino acid sequence of the human secreted protein AK647.				
KM	AK647; aortic tissue development; smooth muscle cell modulator; SCID;				
KM	nutritional supplement; vasculogenesis; embryonic development; infection;				
KM	cytokine activity; cell proliferation; cell differentiation; detect; HIV;				
KM	immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;				
KM	wound healing; restenosis; atherosclerosis; drug screen.				
OS	Homo sapiens.				
PN	W09960125-A2.				
PD	25-NOV-1999.				
PF	18-MAY-1999;	99WO-US10931.			
PR	19-MAY-1998;	98US-0081002.			
PR	21-MAY-1998;	98US-0083002.			
PA	(GENY) GENETICS INST INC.				
PI	Jacobs K, McCoy JM, Racie L,	LaVallie E, Treacy M, Evans C;			
PI	Agostino M, Lu Z, Merberg D;				
WP	WPI: 2000-053298/04.				
DR	N-PSDB: AAZ39892.				
XX	Proteins, and their encoding polynucleotides, used for treating e.g.				

FT Region /Label- EGF-4
 FT 272..314 /Label- EGF-5
 XX MO9846746-A1. (555)
 XX 22-OCT-1998.
 XX 11-APR-1997; 97MO-US06020.
 XX 11-APR-1997; 97MO-US06020.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX L1 H, Olsen HS;
 XX WPI: 1998-568728/48.
 XX N-PSDB: AAV62432.
 DR New isolated extracellular/epidermal growth factor - used for
 PT regulating vascular smooth muscle cell proliferation, e.g. for
 DR enhancing neurological functions or treating neoplasia and other
 XX disorders.
 XX Claim 10a; Fig 1A-D; 62pp; English.

This sequence represents a novel human extracellular/epidermal growth factor-like protein, EGF. This protein can be used to regulate vascular smooth muscle cell proliferation and for restoration or enhancement of neurological functions diminished as a result of other damaging pathologies such as AIDS dementia. The protein can also be used to treat senile dementia, ocular disorders such as corneal inflammation, for targeting tumour cells, for treating kidney disorders, for liver regeneration or treating liver dysfunction, for treating wounds including all cutaneous wounds, corneal wounds, and injuries to the epithelial-lined hollow organs of the body or resulting from trauma such as burns, abrasions and cuts as well as from surgical procedures such as surgical incisions and skin grafting. The polypeptides can also be used for treating chronic conditions, such as chronic ulcers, diabetic ulcers, other non-healing (trophic) conditions, to treat Marfan syndrome, to promote hair follicular development, to stimulate growth and differentiation of various epidermal and epithelial cells in vivo and in vitro and to stimulate embryogenesis. Antagonists to EGF can be used to treat neoplasia such as cancers or tumours, skin disorders such as psoriasis or corneal inflammation. The products can also be used for identifying EGF receptors, detection, diagnosis and drug screening.

Sequence 448 AA:

Query Match 100.0%; Score 2533; DB 19; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPGKRLITVTITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTTPEACRGDMCV 60
 1 MPGKIRILVTITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTTPEACRGDMCV 60
 61 NONGGILCIPRTNPPVYRGPNSTPYSGPYPAAPLSPNPTIRPLICRGYOMD 120
 61 NONGGILCIPRTNPPVYRGPNSTPYSGPYPAAPLSPNPTIRPLICRGYOMD 120
 61 NONGGILCIPRTNPPVYRGPNSTPYSGPYPAAPLSPNPTIRPLICRGYOMD 120
 121 ESNQCVYVDECAVDSHOQCNPTQICINTEGGYCTSDGTYMLLEGCLDIDECRYGCOOL 180
 121 ESNQCVYVDECAVDSHOQCNPTQICINTEGGYCTSDGTYMLLEGCLDIDECRYGCOOL 180
 181 CANVPSSYSTCTCNPGLTINEDGRSCQDVNECATENPCVQTGVNTYGSYICRDPGELEE 240
 181 CANVPSSYSTCTCNPGLTINEDGRSCQDVNECATENPCVQTGVNTYGSYICRDPGELEE 240
 241 DGVHCSDMDECSSEFLCQHECVNQPGTYFSCPGPYTLLDNRSCQDINCECHNHTCN 300
 241 DGVHCSDMDECSSEFLCQHECVNQPGTYFSCPGPYTLLDNRSCQDINCECHNHTCN 300

QY 301 LQOTCYNLGGFKCIDPICEEPLRLISDNCKMCPAENPGCRDQPFLLYRDMVYSGRS 360
 DB 301 LQOTCYNLGGFKCIDPICEEPLRLISDNCKMCPAENPGCRDQPFLLYRDMVYSGRS 360
 QY 361 VPADIFOMQATTRPGAYYIFQIKSGNKGREFFYAROTGPISATLVTRPIKPREQLDL 420
 DB 361 VPADIFOMQATTRPGAYYIFQIKSGNKGREFFYAROTGPISATLVTRPIKPREQLDL 420
 QY 421 EMITVNTVINEGSSVTRLRIVYSQYPP 448
 DB 421 EMITVNTVINEGSSVTRLRIVYSQYPP 448

RESULT 2

AAW95709
 ID AAW95709 standard; protein: 448 AA.
 AC AAW95709;
 DT 21-JUN-1999 (first entry)
 XX Homo sapiens fetal kidney clone AK647 secreted protein.
 XX Secreted protein: fetal kidney.
 XX Homo sapiens.
 XX MO9900405-A1. (555)
 XX 07-JAN-1999.
 XX 29-JUN-1998; 98MO-US13530.
 XX 30-JUN-1997; 97US-0885610.
 XX (GEMT) GENETICS INST INC.
 XX Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 PI Merberg D, Racle LA, Treacy M;
 DR WPI: 1998-095671/08.
 DR N-PSDB: AAX07567.
 XX New polynucleotides encoding secreted human proteins - are derived
 PT from foetal kidney or adult retina cDNA libraries, used as, e.g.
 PT potential vaccines
 XX Claim 11; Pages 52-54; 76pp; English.

The sequence is that of a secreted protein from a human fetal kidney clone AK296. Such a sequence is predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, CC haemopoiesis regulating activity, tissue growth activity, CC activin/inhibin regulating activity, chemotactic/chemokinetic activity, CC activin/thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. It is also stated to be useful for gene therapy.

Sequence 448 AA:

Query Match 100.0%; Score 2533; DB 20; Length 448;
 Best Local Similarity 100.0%; Pred. No. 3.2e-164;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPGKRLITVTITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTTPEACRGDMCV 60
 1 MPGKIRILVTITIALCLPSPGNAOCTNGFDLDRSGQCCLDIDECRTTPEACRGDMCV 60


```
OY 1 MEGIKRILVVTIILALCLPSPGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMCV 60
DB 1 MEGIKRILVVTIILALCLPSPGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMCV 60
OY 61 NONGGVLICIPRTNPVYRGPSNPYSTPSGYPAPAAPLSAPNYPITISRPLICRGYQMD 120
DB 61 NONGGVLICIPRTNPVYRGPSNPYSTPSGYPAPAAPLSAPNYPITISRPLICRGYQMD 120
OY 121 ESNQCVDEVCANVDSHQCNPYQICINTEGTYCTSDGYWLLGQCLDIDECRYGCOOL 180
DB 121 ESNQCVDEVCANVDSHQCNPYQICINTEGTYCTSDGYWLLGQCLDIDECRYGCOOL 180
OY 181 CANVPGSYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGYELE 240
DB 181 CANVPGSYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGYELE 240
OY 241 DGVHCSMDMECSFSEFLCQHCVNQPGTYFCSCPFGYILLDDNRSCQDINECEHRNHTCN 300
DB 241 DGVHCSMDMECSFSEFLCQHCVNQPGTYFCSCPFGYILLDDNRSCQDINECEHRNHTCN 300
OY 301 LOOTCYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDOFTILYRDMDVYSGRS 360
DB 301 LOOTCYNLOGGFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDOFTILYRDMDVYSGRS 360
OY 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQGTGPISATLVMTRPPIKGPRIQOLDL 420
DB 361 VPADIFOMQATTRYPGAYYIFQIKSGNEGREFYMRQGTGPISATLVMTRPPIKGPRIQOLDL 420
OY 421 EMITVNTVINFRGSSVYRLRIYVSQYPF 448
DB 421 EMITVNTVINFRGSSVYRLRIYVSQYPF 448
```

Search completed: July 3, 2003, 18:26:16
Job time : 89.4684 secs

;; CURRENT APPLICATION NUMBER: US/10/176,757
;; CURRENT FILING DATE: 2002-06-20
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-176-757-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITALLCPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
DB 1 MGKIRILVTITALLCPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
Y 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCVYDDECATDSHOCNPTQICINTEGGYTCSTDGWLLLEGCLDIDECRYGCOOL 180
DB 121 ESNOCVYDDECATDSHOCNPTQICINTEGGYTCSTDGWLLLEGCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGELEE 240
DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGELEE 240
QY 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDDNRSCODINECEHRNHTCN 300
DB 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDDNRSCODINECEHRNHTCN 300
QY 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
DB 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
DB 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448

RESULT 14

US-10-176-913-408
;; Sequence 408, Application US/10176913
;; Publication No. US20030022298A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C66
;; CURRENT APPLICATION NUMBER: US/10/176,913
;; CURRENT FILING DATE: 2002-06-20
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT

;; ORGANISM: Homo Sapien
US-10-176-913-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITALLCPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
DB 1 MGKIRILVTITALLCPSPGNAOACTNGFDLDROSGCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPNVYRGPNYSPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCVYDDECATDSHOCNPTQICINTEGGYTCSTDGWLLLEGCLDIDECRYGCOOL 180
DB 121 ESNOCVYDDECATDSHOCNPTQICINTEGGYTCSTDGWLLLEGCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGELEE 240
DB 181 CANVPGSYCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGELEE 240
QY 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDDNRSCODINECEHRNHTCN 300
DB 241 DGVHCSMDCECSFSEFLQHECVNQPCTYFCSCPBGYILLDDNRSCODINECEHRNHTCN 300
QY 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
DB 301 LOOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCAENPGCDOPFTILYRMDVYSGRS 360
QY 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
DB 361 VPADIFOMQATTRPGAYYIFQIKSGNEGREFYMRQTPISATLVMPRIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448

RESULT 15

US-10-180-552-408
;; Sequence 408, Application US/10180552
;; Publication No. US20030022300A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C153
;; CURRENT APPLICATION NUMBER: US/10/180,552
;; CURRENT FILING DATE: 2002-06-25
;; Prior Application removed - See file Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 408
;; LENGTH: 448
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-180-552-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; US-10-175-752-408

```

```

Query Match      99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSQCCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSQCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYDVCATDSHOCNPTQICINTEGGYTCSTDYWLLEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVYDVCATDSHOCNPTQICINTEGGYTCSTDYWLLEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
DB 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
QY 241 DGVCSDMDCESEFELQCHQECVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
DB 241 DGVCSDMDCESEFELQCHQECVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
QY 301 LOOTCYNLGGFKCIDPRICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
DB 301 LOOTCYNLGGFKCIDPRICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
QY 361 VPADIFQMOATTRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMTPIKGPRIQDL 420
DB 361 VPADIFQMOATTRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMTPIKGPRIQDL 420
QY 421 EMITVNTVINFRGSSVIRLRITYSOYRF 448
DB 421 EMITVNTVINFRGSSVIRLRITYSOYRF 448

```

```

RESULT 12
US-10-176-482-408
; Sequence 408, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

```

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRP
; ORGANISM: Homo Sapien
; US-10-176-482-408

```

```

Query Match      99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSQCCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRILVTYITIALCLPSPGNAOACTNGFDLDROSQCCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVYRGPSNPSTYSGPYPAAPLAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVYDVCATDSHOCNPTQICINTEGGYTCSTDYWLLEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVYDVCATDSHOCNPTQICINTEGGYTCSTDYWLLEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
DB 181 CANVPGSYCTCNGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
QY 241 DGVCSDMDCESEFELQCHQECVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
DB 241 DGVCSDMDCESEFELQCHQECVNOPTGYFCSCPFGYILLDDNRSCQDINECEHNRHTCN 300
QY 301 LOOTCYNLGGFKCIDPRICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
DB 301 LOOTCYNLGGFKCIDPRICEEPYLRISDNRCMCPAENPGCRDQFTILYRMDVYSGRS 360
QY 361 VPADIFQMOATTRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMTPIKGPRIQDL 420
DB 361 VPADIFQMOATTRYPGAYYIFQIKSGNEGREFYMRQGPISATLVMTPIKGPRIQDL 420
QY 421 EMITVNTVINFRGSSVIRLRITYSOYRF 448
DB 421 EMITVNTVINFRGSSVIRLRITYSOYRF 448

```

```

RESULT 13
US-10-176-757-408
; Sequence 408, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86

```

QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 9

US-10-173-706-408
; Sequence 408, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173.706
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-408

Query Match

Best Local Similarity 99.8%; Score 2527; DB 9; Length 448;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITLALCLPSPGNAQAQCTNGFDLDROSGQCLDIDECRTIPEACRGDMCV 60
Db 1 MGKIRILVTITLALCLPSPGNAQAQCTNGFDLDROSGQCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPYRGFYSNPYSTPGYPAAAPPLSAPNPTISRLICRFQYOMD 120
Db 61 NONGGYLCIPRTNPYRGFYSNPYSTPGYPAAAPPLSAPNPTISRLICRFQYOMD 120
QY 121 ESNOCVADVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLEGGQCLDIDECRYGYCOOL 180
Db 121 ESNOCVADVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLEGGQCLDIDECRYGYCOOL 180
QY 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
Db 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSPGTYILLDNRSCODINECEHRNHTCN 300
Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSPGTYILLDNRSCODINECEHRNHTCN 300
QY 301 LGQTCYNLQGGFKCIDPTRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
Db 301 LGQTCYNLQGGFKCIDPTRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 10

US-10-175-738-408
; Sequence 408, Application US/10175738
; Publication No. US20030022294A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175.738
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-408

Query Match

Best Local Similarity 99.8%; Score 2527; DB 9; Length 448;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGKIRILVTITLALCLPSPGNAQAQCTNGFDLDROSGQCLDIDECRTIPEACRGDMCV 60
Db 1 MGKIRILVTITLALCLPSPGNAQAQCTNGFDLDROSGQCLDIDECRTIPEACRGDMCV 60
QY 61 NONGGYLCIPRTNPYRGFYSNPYSTPGYPAAAPPLSAPNPTISRLICRFQYOMD 120
Db 61 NONGGYLCIPRTNPYRGFYSNPYSTPGYPAAAPPLSAPNPTISRLICRFQYOMD 120
QY 121 ESNOCVADVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLEGGQCLDIDECRYGYCOOL 180
Db 121 ESNOCVADVDECATDHSQCNPQTQICINTEGGYTCSCTDGYWLEGGQCLDIDECRYGYCOOL 180
QY 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
Db 181 CANVGSYSCTCNPFTLNEDGRSCODVNECATENPCVOTCVNTYGSFICRDPGYLEE 240
QY 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSPGTYILLDNRSCODINECEHRNHTCN 300
Db 241 DGVHCSMDDECSFSEFLCOHECVNPGTYFCSPGTYILLDNRSCODINECEHRNHTCN 300
QY 301 LGQTCYNLQGGFKCIDPTRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
Db 301 LGQTCYNLQGGFKCIDPTRCEEPYLRISDNRCMCPAENPGCRDQPTIILYRDMDVYSGRS 360
QY 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
Db 361 VPADIFOMQATRRYPGAYYIFQIKSGNREGREFYMQTGPISATLVWTRPIKGPRIQOLDL 420
QY 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448
Db 421 EMITVNTVINRGSSVIRLRIVYSQYPF 448

RESULT 11

US-10-175-752-408
; Sequence 408, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:

```
Db 121 ESNQCVYDDECATDHSQCNPFOICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Qy 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Qy 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
Qy 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Qy 361 VPADIFOMQATTTPGAYYIFQIKSGNEGREFFYRQTPISATLVMPRIKGPRIQLDL 420
Db 361 VPADIFOMQATTTPGAYYIFQIKSGNEGREFFYRQTPISATLVMPRIKGPRIQLDL 420
Qy 421 EMITVNTVINFRSSVIRLRIRIYVSQYPF 448
Db 421 EMITVNTVINFRSSVIRLRIRIYVSQYPF 448
```

RESULT 7

```
US-10-176-758-408
; Sequence 408, Application US/10176758
; Publication No. US2003008353A1
```

```
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
; -10-176-758-408
```

```
Query Match 99.88; Score 2527; DB 9; Length 448;
Best Local Similarity 99.88; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MPGIKRLITVTITIALCLPSGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMVCV 60
Db 1 MPGIKRLITVTITIALCLPSGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMVCV 60
Qy 61 NONGGYLCIRPTNPVYRGPSNPYSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Db 61 NONGGYLCIRPTNPVYRGPSNPYSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Qy 121 ESNQCVYDDECATDHSQCNPFOICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Db 121 ESNQCVYDDECATDHSQCNPFOICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Qy 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Qy 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
```

```
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
Qy 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Qy 361 VPADIFOMQATTTPGAYYIFQIKSGNEGREFFYRQTPISATLVMPRIKGPRIQLDL 420
Db 361 VPADIFOMQATTTPGAYYIFQIKSGNEGREFFYRQTPISATLVMPRIKGPRIQLDL 420
Qy 421 EMITVNTVINFRSSVIRLRIRIYVSQYPF 448
Db 421 EMITVNTVINFRSSVIRLRIRIYVSQYPF 448
```

RESULT 8

```
US-10-175-737-408
; Sequence 408, Application US/10175737
; Publication No. US20030013153A1
```

```
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C150
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
; -10-175-737-408
```

```
Query Match 99.88; Score 2527; DB 9; Length 448;
Best Local Similarity 99.88; Pred. No. 2.1e-164;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 MPGIKRLITVTITIALCLPSGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMVCV 60
Db 1 MPGIKRLITVTITIALCLPSGNAQAQCTNGFDLDRSGQCLDIDECRTIPEACRGDMVCV 60
Qy 61 NONGGYLCIRPTNPVYRGPSNPYSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Db 61 NONGGYLCIRPTNPVYRGPSNPYSTPYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
Qy 121 ESNQCVYDDECATDHSQCNPFOICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Db 121 ESNQCVYDDECATDHSQCNPFOICINTEGGYTCSTGDTGWLLEGGCLDIDECRYGCAQL 180
Qy 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Db 181 CANVPGSYSTCTCNPGLFNLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRCDPGELEE 240
Qy 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
Db 241 DGVHCSMDMECSSEFLCOHECVNQPGTYFCSCPFGYILLDDNRSQDINCEHNRHNCN 300
Qy 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
Db 301 LOOTCYNLOGGFECIDPICEEPLRISDNRCMCPAENPGCRDOPFTILYRDMDVSGRS 360
```

```

; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403296
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/423844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/522342
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: 09/548815
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 09/664610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/767609
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/808689
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/866028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/870574
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/886342
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: PCT/US98/14552
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US98/18824
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/19093
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: PCT/US98/25190
; PRIOR FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

```

Query Match 99.8%; Score 2527; DB 9; Length 448;

Best Local Similarity 99.8%; Pred. No. 2,1e-164; Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MFGIRILVTTLALCLPSPGNAOACTNGFDDDRSGGCLDIDECRTIPACRGDMCV 60
DB 1 MFGIRILVTTLALCLPSPGNAOACTNGFDDDRSGGCLDIDECRTIPACRGDMCV 60

```

```

QY 61 NONGGYLCIPRTNPVRYSGPNPSTPYSGPYAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVRYSGPNPSTPYSGPYAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCYDVDECATDHSQCNPQTQICINTEGGVTCSCDTGWMLEGCCLDIDECRRGYCOOL 180
DB 121 ESNOCYDVDECATDHSQCNPQTQICINTEGGVTCSCDTGWMLEGCCLDIDECRRGYCOOL 180
QY 181 CANVPGSYSCCTCPNPFTLNEDGSRSCQDVNCAETENPCVQTCVNTYGSFICRCDBGYELE 240
DB 181 CANVPGSYSCCTCPNPFTLNEDGSRSCQDVNCAETENPCVQTCVNTYGSFICRCDBGYELE 240
QY 241 DGVHCSMDCECSSEFLCOHECVNQPCTYSCSPGTYLLDDNNSCODINCEHRNHTCN 300
DB 241 DGVHCSMDCECSSEFLCOHECVNQPCTYSCSPGTYLLDDNNSCODINCEHRNHTCN 300
QY 301 LOOTCYNLOGGFCIPDIRCEEPYLRISDRNCMPAENPCGRDQPTILYRDMOVSGRS 360
DB 301 LOOTCYNLOGGFCIPDIRCEEPYLRISDRNCMPAENPCGRDQPTILYRDMOVSGRS 360
QY 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFTWROTGPISATLVMTRPPIKPREIOLDL 420
DB 361 VPADIFQOMATTRYPGAYYIFQIKSGNEGREFTWROTGPISATLVMTRPPIKPREIOLDL 420
QY 421 EMITVTVINFRGSSVIRLRITYSOYF 448
DB 421 EMITVTVINFRGSSVIRLRITYSOYF 448

```

RESULT 6
US-10-174-590-408

; Sequence 408, Application US/10174590
; Publication No. US20030008352A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C42

; CURRENT APPLICATION NUMBER: US/10/174,590

; CURRENT FILING DATE: 2002-06-18

; Prior application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 408

; LENGTH: 448

; TYPE: PRT

; ORGANISM: Homo Sapien

; US-10-174-590-408

Query Match 99.8%; Score 2527; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,1e-164; Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MFGIRILVTTLALCLPSPGNAOACTNGFDDDRSGGCLDIDECRTIPACRGDMCV 60
DB 1 MFGIRILVTTLALCLPSPGNAOACTNGFDDDRSGGCLDIDECRTIPACRGDMCV 60
QY 61 NONGGYLCIPRTNPVRYSGPNPSTPYSGPYAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGGYLCIPRTNPVRYSGPNPSTPYSGPYAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNOCYDVDECATDHSQCNPQTQICINTEGGVTCSCDTGWMLEGCCLDIDECRRGYCOOL 180
DB 121 ESNOCYDVDECATDHSQCNPQTQICINTEGGVTCSCDTGWMLEGCCLDIDECRRGYCOOL 180

```


Oy 361 VPADIFQMATTTRYGAYVYIFQIKSGNCGREFYMGOTGPISATLVMTRPKSPREIOLDL 420
Db 361 VPADIFQMATTTRYGAYVYIFQIKSGNCGREFYMGOTGPISATLVMTRPKSPREIOLDL 420
Oy 421 EMITVTVINFRGSSVIRLRIYVSQYRF 448
Db 421 EMITVTVINFRGSSVIRLRIYVSQYRF 448

RESULT 5
US-10-066-500-15
; Sequence 15, Application US/10066500
; Patent No. US20020177165A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Bolstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Eaton
; APPLICANT: Napoleone Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gertlisen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kjaevlin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P130R1C7
; CURRENT APPLICATION NUMBER: US/10/066,500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092

; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106032
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/109304
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/139695
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/145070
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/149396
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 08/918874
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 08/933821
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 08/960507
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 09/114844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: 09/136801
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136804
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136828
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/158342
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 09/202088
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/254460
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/254465
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 09/284663
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 09/332928
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/332929
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333075
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333077
; PRIOR FILING DATE: 1999-06-14

STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/839,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF224
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8509
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-275-805-2

Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
OY 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPITSRPLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPITSRPLICRFGYOMD 120
OY 121 ESNOCVDDDECATSDHOCNPQICINTEGGYTCSCDGYWLLGQCLDIDECRGYCOOL 180
DB 121 ESNOCVDDDECATSDHOCNPQICINTEGGYTCSCDGYWLLGQCLDIDECRGYCOOL 180
OY 181 CANPGSYSCNPGFTLNEDRSCQDVNECATENPCVOTCVMTYGSFICRCPGYELE 240
DB 181 CANPGSYSCNPGFTLNEDRSCQDVNECATENPCVOTCVMTYGSFICRCPGYELE 240
OY 241 DGVHCSMDDECSFSEFLQHECVNOPGTYFCSCPFGYILLDNRSCODINECHRNHTCN 300
DB 241 DGVHCSMDDECSFSEFLQHECVNOPGTYFCSCPFGYILLDNRSCODINECHRNHTCN 300
OY 301 LQOTCYNLOGGFKCIDPRICEEPYLRISNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360
DB 301 LQOTCYNLOGGFKCIDPRICEEPYLRISNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360
OY 361 VPAIFQMAQATTRYGAYYIFQIKSGNEGREFYMRGTGPISATLVMTRP1KGRREIOLDL 420
DB 361 VPAIFQMAQATTRYGAYYIFQIKSGNEGREFYMRGTGPISATLVMTRP1KGRREIOLDL 420
OY 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIYVSOYPF 448

RESULT 4
US-09-836-561-1
; Sequence 1, Application US/09836561
; Patent No. US20020038006A1

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/836,561
FILING DATE: 16-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/212,168
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-836-561-1
Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
DB 1 MPGIRLITVTLALCLPSGNAQAQCTNGFDLRSGQCLDIDECRTIPEACRGDMCV 60
OY 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPITSRPLICRFGYOMD 120
DB 61 NONGGYLCIPRTNPYRGPSNPYSTPYSGPYPAAPPLSAPNYPITSRPLICRFGYOMD 120
OY 121 ESNOCVDDDECATSDHOCNPQICINTEGGYTCSCDGYWLLGQCLDIDECRGYCOOL 180
DB 121 ESNOCVDDDECATSDHOCNPQICINTEGGYTCSCDGYWLLGQCLDIDECRGYCOOL 180
OY 181 CANPGSYSCNPGFTLNEDRSCQDVNECATENPCVOTCVMTYGSFICRCPGYELE 240
DB 181 CANPGSYSCNPGFTLNEDRSCQDVNECATENPCVOTCVMTYGSFICRCPGYELE 240
OY 241 DGVHCSMDDECSFSEFLQHECVNOPGTYFCSCPFGYILLDNRSCODINECHRNHTCN 300
DB 241 DGVHCSMDDECSFSEFLQHECVNOPGTYFCSCPFGYILLDNRSCODINECHRNHTCN 300
OY 301 LQOTCYNLOGGFKCIDPRICEEPYLRISNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360
DB 301 LQOTCYNLOGGFKCIDPRICEEPYLRISNRCMCPAENPGCRDOPFTILYRDMDVYSGRS 360

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20020165151A1 Relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-041-016-2

Query Match 100.0%; Score 2533; DB 9; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIRLLVTITIALCLPSPGNMAOCTNGDLDROSQCLDIDECRTIPKCRDMMCV 60
DB 1 MGKIRLLVTITIALCLPSPGNMAOCTNGDLDROSQCLDIDECRTIPKCRDMMCV 60
QY 61 NONGYLICIPRTNPVYRGPSNPNSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPVYRGPSNPNSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVADVDECATDSHQCNPTQICINTEGGYTCSDTGWLLBEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVADVDECATDSHQCNPTQICINTEGGYTCSDTGWLLBEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPNGFTLNEDGRSCDYNECATENPCVQTCVNTYGSITCRDPEYLEE 240
DB 181 CANVPGSYCTCNPNGFTLNEDGRSCDYNECATENPCVQTCVNTYGSITCRDPEYLEE 240
QY 241 DGVHCSMDCECSFEFLCOHECVNQPSTYSCPPGYILLDDNRSCODINECEHNNHCN 300
DB 241 DGVHCSMDCECSFEFLCOHECVNQPSTYSCPPGYILLDDNRSCODINECEHNNHCN 300
QY 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
DB 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
QY 361 VPADIFOMATTRPGAYYIFQIKSGNEGREFYKROTGPISATLVMTPIKGPRIQIDL 420
DB 361 VPADIFOMATTRPGAYYIFQIKSGNEGREFYKROTGPISATLVMTPIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 2
US-09-083-002-2
Sequence 2, Application US/09083002
Patent No. US2001001650A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Racile, Lisa A.
APPLICANT: Lavallee, Edward R.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Agostino, Michael
APPLICANT: Lu, Zhijian
APPLICANT: Honjo, Tasuku
APPLICANT: Tashiro, Kei
APPLICANT: Nakamura, Tomoyuki
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/083.002
FILING DATE:
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-083-002-2

Query Match 100.0%; Score 2533; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.3e-165;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKIRLLVTITIALCLPSPGNMAOCTNGDLDROSQCLDIDECRTIPKCRDMMCV 60
DB 1 MGKIRLLVTITIALCLPSPGNMAOCTNGDLDROSQCLDIDECRTIPKCRDMMCV 60
QY 61 NONGYLICIPRTNPVYRGPSNPNSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
DB 61 NONGYLICIPRTNPVYRGPSNPNSTYSGPYPAAPPLSAPNPTISRPLICRGYQMD 120
QY 121 ESNQCVADVDECATDSHQCNPTQICINTEGGYTCSDTGWLLBEGQCLDIDECRYGCOOL 180
DB 121 ESNQCVADVDECATDSHQCNPTQICINTEGGYTCSDTGWLLBEGQCLDIDECRYGCOOL 180
QY 181 CANVPGSYCTCNPNGFTLNEDGRSCDYNECATENPCVQTCVNTYGSITCRDPEYLEE 240
DB 181 CANVPGSYCTCNPNGFTLNEDGRSCDYNECATENPCVQTCVNTYGSITCRDPEYLEE 240
QY 241 DGVHCSMDCECSFEFLCOHECVNQPSTYSCPPGYILLDDNRSCODINECEHNNHCN 300
DB 241 DGVHCSMDCECSFEFLCOHECVNQPSTYSCPPGYILLDDNRSCODINECEHNNHCN 300
QY 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
DB 301 LQOTCYNLOGGFKCIDIRCEEPYLRISDNRCMCPAENPGCRDQFTLLYRDMDVSGRS 360
QY 361 VPADIFOMATTRPGAYYIFQIKSGNEGREFYKROTGPISATLVMTPIKGPRIQIDL 420
DB 361 VPADIFOMATTRPGAYYIFQIKSGNEGREFYKROTGPISATLVMTPIKGPRIQIDL 420
QY 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448
DB 421 EMITVNTVINFRGSSVIRLRIVYSQYPF 448

RESULT 3
US-09-275-805-2

Sequence 2, Application US/09275805
Patent No. US20010051358A1

GENERAL INFORMATION:

APPLICANT: OLSEN, HENRIK S.
APPLICANT: LI, HAODONG
TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
TITLE OF INVENTION: LIKE PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:53:42 ; Search time 88.4684 Seconds

(Without alignments)
582.319 Million cell updates/sec

Title: US-09-674-379a-13

Perfect score: 2533
Sequence: 1 MGKIRLITLTLALCPSP.....INRGSSVIRLRIVSQYRF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications-AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2533	100.0	448	9	US-10-041-016-2
2	2533	100.0	448	10	US-09-083-002-2
3	2533	100.0	448	10	US-09-275-805-2
4	2533	100.0	448	10	US-09-836-561-1
5	2527	99.8	448	9	US-10-066-500-15
6	2527	99.8	448	9	US-10-174-590-408
7	2527	99.8	448	9	US-10-176-758-408
8	2527	99.8	448	9	US-10-175-737-408
9	2527	99.8	448	9	US-10-173-706-408
10	2527	99.8	448	9	US-10-173-738-408
11	2527	99.8	448	9	US-10-175-752-408
12	2527	99.8	448	9	US-10-176-482-408
13	2527	99.8	448	9	US-10-176-757-408
14	2527	99.8	448	9	US-10-176-913-408
15	2527	99.8	448	9	US-10-180-552-408
16	2527	99.8	448	9	US-10-180-557-408
17	2527	99.8	448	9	US-10-173-700-408
18	2527	99.8	448	9	US-10-174-572-408
19	2527	99.8	448	9	US-10-174-579-408

20	2527	99.8	448	9	US-10-174-582-408	Sequence 408, App
21	2527	99.8	448	9	US-10-174-588-408	Sequence 408, App
22	2527	99.8	448	9	US-10-175-739-408	Sequence 408, App
23	2527	99.8	448	9	US-10-175-740-408	Sequence 408, App
24	2527	99.8	448	9	US-10-175-743-408	Sequence 408, App
25	2527	99.8	448	9	US-10-176-488-408	Sequence 408, App
26	2527	99.8	448	9	US-10-176-492-408	Sequence 408, App
27	2527	99.8	448	9	US-10-176-747-408	Sequence 408, App
28	2527	99.8	448	9	US-10-176-750-408	Sequence 408, App
29	2527	99.8	448	9	US-10-176-985-408	Sequence 408, App
30	2527	99.8	448	9	US-10-176-987-408	Sequence 408, App
31	2527	99.8	448	9	US-10-176-991-408	Sequence 408, App
32	2527	99.8	448	9	US-10-176-992-408	Sequence 408, App
33	2527	99.8	448	9	US-10-176-993-408	Sequence 408, App
34	2527	99.8	448	9	US-10-184-658-408	Sequence 408, App
35	2527	99.8	448	9	US-10-002-796-15	Sequence 15, App1
36	2527	99.8	448	9	US-10-066-273-15	Sequence 15, App1
37	2527	99.8	448	9	US-10-066-494-15	Sequence 15, App1
38	2527	99.8	448	9	US-10-173-695-408	Sequence 408, App
39	2527	99.8	448	9	US-10-173-697-408	Sequence 408, App
40	2527	99.8	448	9	US-10-173-705-408	Sequence 408, App
41	2527	99.8	448	9	US-10-174-576-408	Sequence 408, App
42	2527	99.8	448	9	US-10-174-585-408	Sequence 408, App
43	2527	99.8	448	9	US-10-174-586-408	Sequence 408, App
44	2527	99.8	448	9	US-10-175-747-408	Sequence 408, App
45	2527	99.8	448	9	US-10-176-481-408	Sequence 408, App

ALIGNMENTS

RESULT 1
US-10-041-016-2
; Sequence 2, Application US/10041016
; Patent No. US20020165151A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racie, Lisa A.
Lavallee, Edward R.
Merberg, David
Trecay, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasuku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851



1
2

Db 663 MTORGCEIDDECLNPS-TC-PDEQCVNPSGYCPCBTBGRMGNGOCLDYDECLPNV 720
QY 177 CQO-LEANNPGSISCTCNPGFTLNEGRSCQDVNECAPCVQ----- 219
Db 721 CANGCDNLEGSYMCCHKGYTRTPDHKCRDIDECQOGLCVNGQCKNTGSEFRCTGQ 780
QY 220 -----TCVNTYGSFICRCDPGYLEEDGVHCSMDCECF 253
Db 781 GYOLSAKQCEIDECQHRHLCAHGOCNRTGSEFQVCDQGYRASGLGDHCEIDINECLE 840
QY 254 SEFLCOH-BCVNPQTYFSCPCPGYILLDDNRSCODINECEHNRHTCNLOQTCYNLOGCF 312
Db 841 DKSVCRGDCINTAGSYDCTCPDGF-QLDDNKTQDINECEHPG-LGGPGCECLNTGSEF 898
QY 313 KCI-----DPRCEPYLRIS-----DN-----RCMC-----PAENPGCR 342
Db 899 HCVCOQGSISADGRTCEIDECVNTVCDSHGFCNTAGSFRLCYGQFAPDQGGCV 958
Y 343 DQFTILYRDMDVSG 358
Db 959 D-----VNECELLSG 968

RESULT 12
US-08-897-443-1
Sequence 1, Application US/08897443
Patent No. 5981263

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Mathew
TITLE OF INVENTION: HUMAN MATRILIN-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,443
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PE-0348 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: UTRSN0702
CLONE: 681719
US-08-897-443-1

Query Match 18.8%; Score 475; DB 2; Length 638;
Best Local Similarity 28.4%; Pred. No. 5,4e-30;

Matches 96; Conservative 61; Mismatches 117; Indels 64; Gaps 11;
QY 15 LCLPSPGNAOCTNEDFDLROSGOCLDDECETIPACGDMACVONGICYLICPTNP 74
Db 293 LCVNPPGSSVCCYCYSTALAEDEGRCVADVYCASENHGCHE-CVNADSYLC----- 344
QY 75 YVRGYSNDYSPYSGPYPAAPLAPNPTLSRPLICRFGYQMD-ESNQCVVDCEAT 133
Db 345 -----QCHGEGALNNDKTKTKYIDVCS 367
QY 134 DSHQCNPTQICINTEGGYTCSTGYWLEGO--CLDIDECRYG--CQOLCANVPGSYS 189
Db 368 SNHGQ--HECVNTDYSYCHCLKPTLPDCKTKRINICALNKPGCEHCVMMESY 425
QY 190 CTGNPGFTLNEGRSCQDVNECA-TENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMD 248
Db 426 CRCHRGYITLDPNGKTSRVDHCAQDHDGEBQLCLNTEDSFVCCSSEFTINEDLKTESRV 485
QY 249 DECSFSEFLQHECVNQPEYFSCPCPGYILLDDNRSCODINECEHNRHTCNLOQTCYNL 308
Db 486 DYCLSDHGCSEYSCVMMDSFACQCEGHVLRSDGKTCALDKSCALGDHGC--EHSVSS 543
QY 309 OGFEKCIDPRCEPYLRISDNR-C-----MCPAENPGC 341
Db 544 EDSFVC-----QCFEGITLREDGKTCRRKQCAIDHGC 577

RESULT 13

US-08-479-722B-2
Sequence 2, Application US/08479722B
Patent No. 6074840

GENERAL INFORMATION:

APPLICANT: Bonadio, Jeffrey
APPLICANT: Yin, Mushan
TITLE OF INVENTION: LATENT TGF β BINDING PROTEIN (LTBP)
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Williams, Morgan & Amerson
STREET: 7676 Hillmont, Suite 250
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77040

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,722B
FILING DATE: 07-JUN-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US95/02251
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994

ATTORNEY/AGENT INFORMATION:

NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100,000500/EUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7011
TELEFAX: (713) 934-7011

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 183 amino acids
TYPE: amino acid

```

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 458228
US-09-212-168-5

Query Match
Best Local Similarity 48.3%; Score 963.5; DB 4; Length 387;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

QY 96 APPLSAPNPTTISRPLICRFGYOMDESNOVCVDECATDSHCNPTQICINTEGGYTSC 155
DB 38 ABOQRIIPNR--SHRIQACAGYEGSEHNVCODIDECTAGTINCRCADYICINLRGSFACQC 95
QY 156 TCGYMLLEGCLIDDERY-CYCOQLCANVPGSCTCNPEFTLNEDGRSCQDVNECATE 214
DB 96 PRGYQRGECVVIDECTIPPYCHQRGVNTPGSFQCSPGFOLAAANNVTCVDINECDAS 155
QY 215 NPCVQTCVNTYGSFICRCDPEYLEEDGVHCSMDSECSFSEFLCOHECVNOPGTVEGSC 274
DB 156 NCAQOCYNILGSLFICOCNOGIELSDRLNEDIDECTSSYLCOYOCVNEPGRKSCMP 215
QY 275 PGYILLDNRSQCINDECEHNRHNTCNLQCTCYNLGGFKCIDPICEEPYLRISDNRCMC 334
DB 216 OGQYVY-RSRFCODINECETTNE-CRDEMCNMYHGFRCPYRNPCCOPYILITPENRCVC 273
QY 335 PAENPCGRQPTILYRDMVVSGRSVPADIFQWQATTRYGAYIIFQIKSNEGREFRM 394
DB 274 PVSNAACRELPQSIYKYSIRSDRSVSDIFQIQTATTIVANTINTFRKISGENGEFYL 333
QY 395 ROTGPISATLVTRPIKGPREIQLDLEMTVNTVINFGSSVIRLRIYVSQYPF 448
DB 334 RQTSFVSAMLYVKSLSGPREHIVDLEMLTVSSIGTRFTSSVLRLLITVGPSPF 387

```

```

APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for windows Version 3.0
SEQ ID NO 336
LENGTH: 274
TYPE: PR1
ORGANISM: Human
US-09-188-930-336

Query Match
Best Local Similarity 36.0%; Score 912.5; DB 4; Length 274;
Matches 154; Conservative 47; Mismatches 73; Indels 1; Gaps 1;

QY 174 YGICQQLCANVPGSYCTCNPFILNEDGRSCQDVNECATENPCVQTCVNTYGSFICRD 233
DB 1 YRCQHRQVNLPGSFRCQCEPFGOLGPNRSCVDNEDMGACBQRCFNSYGTFLCRCH 60
QY 234 PGYLEEDGVHCSMDSECSFSEFLCOHECVNOPGTVEGSCPRGYTILLDNRSQCINDECE 293
DB 61 QGYELHNDGFSQSIDECSSYSLCQYRCVNEPGRFSCHCPQGYOLL-ATRLCQDIDECB 119
QY 294 HRNHTCNLQCTCYNLGGFKCIDPICEEPYLRISDNRCMCPAENPCGRDQPTILYRDM 353
DB 120 SGAHQCEAQTCVNFHGRCVDTNRCVEPIYQVSENRCCLCPASNPLCREQSSIVHRYM 179
QY 354 DVYSGRSVPADIFQWQATTRYGAYIIFQIKSNEGREFRMOTGPISATLVTRPIKGP 413
DB 180 TITSERVPADVFIQATSVYGVAFNAFOIRNNSQSGDFYIKQINNVSAMLVLRVTPGP 239
QY 414 RETQLDEMTVNTVINFGSSVIRLRIYVSQYPF 448
DB 240 REYVLDEMTVNTVINFGSSVIRLRIYVSQYPF 274

RESULT 11
5177197-30
Patent No. 5177197
APPLICANT: KANZAKI, TETSUO; OLOFSSON, ANDERS; MOREN, ANITA;
WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,
LENA; HELDIN, CARL-HENRIK
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
NUMBER OF SEQUENCES: 53
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/487,343
FILING DATE: 27-FEB-1990
SEQ ID NO: 30
LENGTH: 1394
5177197-30

Query Match
Best Local Similarity 21.5%; Score 544; DB 6; Length 1394;
Matches 128; Conservative 52; Mismatches 134; Indels 122; Gaps 17;

QY 16 CLSPGNAQOCTGFDLDRSGCLDIDERTYPEACRGDMVCVNONGYLCIPRTNRY 75
DB 562 CINLPVRYTCICIGYFSEQQRKCVVIDECTQVQVHLC-SQRCENTBESFLCI----- 614
QY 76 YRGPEYSNPYSTPYSGPYPAAPPLSAPNPTISRPLI-----CRFGYQ 118
DB 615 -----CPAGFMASEEGTNCIDYDECLRPDYCGEGHCVNTVGAFCRCEYCDSGYR 662
QY 119 MDESNCQVDVDECATDTHSCNPTQICINTEGGYT-SCDYGWMLLEGCLDIDEC-RYGR 176

```

CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,963C
FILING DATE: 11-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NO PCT/US96/05033
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF258
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-963C-9

Query Match 38.0%; Score 963.5; DB 2; Length 387;
Best Local Similarity 48.3%; Pred. No. 5.8e-69;

Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

QY 96 APPLSAPNYPTISRLPLICRFYQMDENOCVVDVDECATSHOCNPTQICINTEGGYTQSC 155
DB 38 ADPQIRIPSNP--SHRIQCAAGYEGSEHNVCODIDECTAGTHNCRAQOVICINLUGSFACQC 95
QY 156 TDGYWLLGEGCLDIDECRY-GYCOOLCANVPGSYSCCTNPGLTNDGRSCODVNECATE 214
DB 96 PPGYQKRGEOQVDIDECTIPYCHQRCVNTPGSFYQCSFGFOLANNNTVCVDINCDAS 155
QY 215 NPCVQTCVNTYGSFICRCDPGYELEEDVHCSMDMDECSFSEFLCOHECVNPGTYFCSCP 274
DB 156 NQCAQOCYNILGSLFCQCNQGYELSSDRINCEIDECRTSSYLCOYQCVNEPGRKFSQKCP 215
QY 275 PGYILLDNRSQODINCEGHRNHTCNLAQTCYNLQGGFKCIDPICEEPLYRLISDNKCMC 334
DB 216 QGYQVY-RSRICQDINECTTNE-CREDEMCMNHYHGFRCPYPRNCPQDPYILTPENRCVC 273
QY 335 PAENGGCNDOPFTILYRDMVSGRSVPADIFOMATTRYPGAYYIFQIKSGNEGRFTM 394
DB 274 PVSNAKRELPQSIYKYKMSIRSDRSVSDIFQIQAATTIYANTINFRKSGNEGEFTL 333
QY 395 ROTGPISATLWTRPIKGRPREIQLDLEMITVNTVINFGRSSVIRLRIYQYRF 448
DB 334 ROTSPVSAMLVLYKSLSGPREHIVDLEMLTVSSIGTFRTSSVLRLLTIYGPFSF 387

RESULT 8

US-08-980-514-3
Sequence 3, Application US/08980514
Patent No. 6004753

GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN SL-5-ECMP-LIKE PROT
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,514
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-845-4166
TELEFAX: 650-845-0555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 458228
US-08-980-514-3

Query Match 38.0%; Score 963.5; DB 3; Length 387;
Best Local Similarity 48.3%; Pred. No. 5.8e-69;

Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

QY 96 APPLSAPNYPTISRLPLICRFYQMDENOCVVDVDECATSHOCNPTQICINTEGGYTQSC 155
DB 38 ADPQIRIPSNP--SHRIQCAAGYEGSEHNVCODIDECTAGTHNCRAQOVICINLUGSFACQC 95
QY 156 TDGYWLLGEGCLDIDECRY-GYCOOLCANVPGSYSCCTNPGLTNDGRSCODVNECATE 214
DB 96 PPGYQKRGEOQVDIDECTIPYCHQRCVNTPGSFYQCSFGFOLANNNTVCVDINCDAS 155
QY 215 NPCVQTCVNTYGSFICRCDPGYELEEDVHCSMDMDECSFSEFLCOHECVNPGTYFCSCP 274
DB 156 NQCAQOCYNILGSLFCQCNQGYELSSDRINCEIDECRTSSYLCOYQCVNEPGRKFSQKCP 215
QY 275 PGYILLDNRSQODINCEGHRNHTCNLAQTCYNLQGGFKCIDPICEEPLYRLISDNKCMC 334
DB 216 QGYQVY-RSRICQDINECTTNE-CREDEMCMNHYHGFRCPYPRNCPQDPYILTPENRCVC 273
QY 335 PAENGGCNDOPFTILYRDMVSGRSVPADIFOMATTRYPGAYYIFQIKSGNEGRFTM 394
DB 274 PVSNAKRELPQSIYKYKMSIRSDRSVSDIFQIQAATTIYANTINFRKSGNEGEFTL 333
QY 395 ROTGPISATLWTRPIKGRPREIQLDLEMITVNTVINFGRSSVIRLRIYQYRF 448
DB 334 ROTSPVSAMLVLYKSLSGPREHIVDLEMLTVSSIGTFRTSSVLRLLTIYGPFSF 387

RESULT 9

US-09-212-168-5
Sequence 5, Application US/09212168
Patent No. 6303765

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

```

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-845-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 443 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRSTN013
CLONE: 2786449
US-08-980-514-1

```

Query Match 50.9%; Score 1289; DB 3; Length 443;

Best Local Similarity 49.9%; Pred. No. 1,1e-94;

Matches 226; Conservative 74; Mismatches 131; Indels 22; Gaps 5;

```

QY 1 MPKIRILVTITIALCLPSPGNAQ-----AQTNGFDLRSGGCLDIDECRTPEACRG 55
DB 8 LPSLLMALLLLLGASAPQDSEPDSTGCTDGYEMDPDSQHCRDNECLTPEACKG 67
QY 56 DMACVNGNGYLCIPRTNPPYRGYSNPSTIPYGAAPPLSAPNPTISRPLICRF 115
DB 68 EMKCMHNGGYLCIPRSAAVINDLHG-----EGP-PPYPFPAOHPPN-----PCPP 111
QY 116 GYOMDNOCVDVDECATDSHOCNPTOICINTEGGYTSCDWYLLGSCCLDIDECRIG 175
DB 112 GYRDPDQSDVDVDECAQALHDCRPSQDCNHLPSGYCTCDEGCKRIKIGECVDIDECRIR 171
QY 176 YCOQLCANVPGSYSCTCNPGFTLINEDGRSCDVNCCATENPCVOTCVNTYGSFICRDPG 235
DB 172 YCHRCVNLPSFRCQCEPFGQLGPNNSCVDNVNCMDGAPCEGCFNSYGFELCRHGG 231
QY 236 YLEBEGVHSDMDECFSEFLCOHECVNOGRTYFCSPPGYILLDDNRSCODINECEHR 295
DB 232 YELHROGFCSDIDECSSYSLCOYRCVNEBGRFSCHPOGQLL-ATRLCODIDECESG 290
QY 296 NHTCNLOTCYNLOGGFKCIDPICEEPYLRISDNRCMPAENPCGROPTILYRDMY 355
DB 291 AHOCSEAGTCVNRHGRICVDYTNRCVEPIYGVSEKCLCPASNPICREPPSSIVHRTMTI 350
QY 356 VSGRVPADIFOMATRRYPGAYYIFQIKSGNEGHEFYMRQTPISATLVMTRPKIGPRE 415
DB 351 TSSRSVPADYFOIATSVYFGAYNAFOIRAGNSOGDFYIROINNVSAMLVLARPYTGPRE 410
QY 416 IQDLEMTYNTYINFRGSSVIRLRIYSOYF 448
DB 411 YVLDLEMTNLSMYSRASSVLRITVEGAYTF 443

```

RESULT 6
US-08-884-072-5

Sequence 5, Application US/08884072

Patent No. 5872234

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl J.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

COUNTY: USA

COUNTY: USA

COUNTY: USA

COUNTY: USA

COUNTY: USA

COUNTY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/884,072

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-845-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 458228

US-08-884-072-5

Query Match 38.0%; Score 963.5; DB 5; Length 387;

Best Local Similarity 48.3%; Pred. No. 5.8e-69;

Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

```

QY 96 APPLSAPNPTISNPLICRFYQNDENOCYVDVDECATDSHOCNPTOICINTEGGYTSCG 155
DB 38 ADPGRIPSNP--SHRIGCAAGYEGSEHNVQDIDECRTAGTNCRADQVCINLRGSAQCC 95
QY 156 TDGTYLLEGGCLDIDECRY-GYCOQLCANVPGSYSCTCNPGFTLINEDGRSCDVNCCATE 214
DB 96 PPGYQKRGEGQVDIDECRTIPYCHQRCVNTGSGSFYCCQSPGQOLANNNTYCDVINCDS 155
QY 215 NPCVQTCVNTYGSFICRCDPEYELLEDGVCSDMDECSFSEFLCOHECVNOGRTYCSGP 274
DB 156 NQCAQCCYNLIGSTICQCNQGYELSSDRLNCEIDECRTSYLCOYQVNEPEKSCMCP 215
QY 275 PGYILLDDNRSCODINECEHNRHTCNLOOTCYNLOGGFKCIDPICEEPYLRISDNRCMC 334
DB 216 QGYQVY-NSRRCQDINEETNE-CREDEMCMNHYGFRCPYRNQCDDPYILPEKRCVC 273
QY 335 PAENPGCRDQFTLLYRDMDVYSGRVPADIFOMATRRYPGAYYIFQIKSGNEGHEFYM 394
DB 274 PVSNAMCRRELPOSIVYKXMSIRSDRSVPSDIFQIATTIYANTINTFRKSSNENGEFYL 333
QY 395 ROTGPISATLVMTRPKIGPREIOLDLEMTYNTVINFGRSSVIRLRIYSOYF 448
DB 334 ROTSPVSAMLVVYSLSGPREHYVLDLEMTVSSITFTRTSSVLRLLTIYGPFSF 387

```

RESULT 7

US-08-833-963C-9

Sequence 9, Application US/08833963C

Patent No. 5916769

GENERAL INFORMATION:

APPLICANT: Olsen, et al.

TITLE OF INVENTION: Extracellular/Epidermal Growth Factor

TITLE OF INVENTION: HCA8A58X

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Ave

CITY: Rockville

STATE: MD

COUNTRY: USA

Matches	448:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps
QY	1	MPGIRKRLITVTITLALCLPSPGNAOACITNGEDLDROSQCCLDIDECRTTPEACRQDMNCV	60					
Db	1	MPGIRKRLITVTITLALCLPSPGMAOACITNGEDLDROSQCCLDIDECRTTPEACRQDMNCV	60					
QY	61	NONGVGLCIPTNTNVRGPGYNPNPYSTPGYPPAAPPLSAPNPPTISRLICRGYOMD	120					
Db	61	NONGVGLCIPTNTNVRGPGYNPNPYSTPGYPPAAPPLSAPNPPTISRLICRGYOMD	120					
QY	121	ESNOCVVDDECATDHSOCNPTQICINTEGGYTCCTDGYWLLGGCLDIDECRYGCOOL	180					
Db	121	ESNOCVVDDECATDHSOCNPTQICINTEGGYTCCTDGYWLLGGCLDIDECRYGCOOL	180					
QY	181	CANPVSYSCTCNGNGFTLINEGRSCODVNECATENPCVOTVNTYGSFRCRCDPGTELEE	240					
Db	181	CANPVSISCTCNGFTLINEGRSCODVNECATENPCVOTVNTYGSFRCRCDPGTELEE	240					
QY	241	DGVHCSMDDECFSFSEFLCOHBCVNOPGTYFPCSPRGYILLDDNRSQCODINECEHNNHCN	300					
Db	241	DGVHCSMDDECFSFSEFLCOHBCVNOPGTYFPCSPRGYILLDDNRSQCODINECEHNNHCN	300					
QY	301	LOQTCVNLQGGFKCIDIPRCEEPYLIRISDNKCMCAENPGCRDOPFTILRYMDVYSGRS	360					
Db	301	LOQTCVNLQGGFKCIDIPRCEEPYLIRISDNKCMCAENPGCRDOPFTILRYMDVYSGRS	360					
QY	361	VPADIFQMOTATRRPCGAYVIRFQIKSGNGREFYMRBOTPISATLVTMRPIKPGPREIQIDL	420					
Db	361	VPADIFQMOTATRRPGAYVIRFQIKSGNGREFYMRBOTPISATLVTMRPIKPGPREIQIDL	420					
QY	421	EMITVNTVINPRGSSVIRLRIYVSQYPR	448					
Db	421	EMITVNTVINPRGSSVIRLRIYVSQYPR	448					

RESULT 2
US-09-212-168-1

Sequence 1, Application US/09212168

Patent No. 6303765

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl J.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/212.168

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/884,072

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PE-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 1:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 448 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: CORNOTO1
: CLONE: 45517
: US-09-212-168-1
:
: Query Match 100.0%; Score 2533; DB 4; Length 448;
: Best Local Similarity 100.0%; Pred. No. 2,6e-193;
: Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0
:
: Oy 1 MPGIRKILTVITLALCLPSFGNAQACNNGFPLDROSQGLDIDECRTPEACRGDMGCV 60
: Db 1 MPGIRKILTVITLALCLPSFGNAQACNNGFPLDROSQGLDIDECRTPEACRGDMGCV 60
: Oy 61 NONGSYLCIPPTNPVYRGPSNPSYSGPYPAAPPLSAPNYPITISRPLICREGYOND 120
: Db 61 NONGSYLCIPPTNPVYRGPSNPSYSGPYPAAPPLSAPNYPITISRPLICREGYOND 120
: Oy 121 ESNOCVDYDECATDSHQCNPTQICINTBEGYTCSCTDGVLWLEGGCLDIDECRYGCOOL 180
: Db 121 ESNOCVDYDECATDSHQCNPTQICINTBEGYTCSCTDGVLWLEGGCLDIDECRYGCOOL 180
: Oy 181 CANVPGSYCTCNPEFTLTNEDRSQDVECATENPCVQTCVNTYGSFICRCDPGYELEB 240
: Db 181 CANVPGSYCTCNPEFTLTNEDRSQDVECATENPCVQTCVNTYGSFICRCDPGYELEB 240
: Oy 241 DGVHCSMDDECSFSFELQHECVNPPGYFCSCPPGYILLDNRSCQDINECHRNHTCN 300
: Db 241 DGVHCSMDDECSFSFELQHECVNPPGYFCSCPPGYILLDNRSCQDINECHRNHTCN 300
: Oy 301 LQOTCYNLOGGKCIDPIRCEEPYLRISDNRCMCAPENPGCRDOPFTIYLRDMDVYSGRS 360
: Db 301 LQOTCYNLOGGKCIDPIRCEEPYLRISDNRCMCAPENPGCRDOPFTIYLRDMDVYSGRS 360
: Oy 361 VPADIFOMQATTRYGAYYIFQIKSGNDEGREFYMRGTGISATLVMTPIKGPRTIOLDL 420
: Db 361 VPADIFOMQATTRYGAYYIFQIKSGNDEGREFYMRGTGISATLVMTPIKGPRTIOLDL 420
: Oy 421 EMITVNTVINFRGSSVIRLRIYVSQYF 448
: Db 421 EMITVNTVINFRGSSVIRLRIYVSQYF 448
:
: RESULT 3
: US-09-188-930-186
: Sequence 186; Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
: APPLICANT: Murison, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 186
: LENGTH: 337
: TYPE: PRT
: ORGANISM: Rat
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (14)...(124)
: NAME/KEY: UNSURE
: LOCATION: (135)...(135)

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:54:52 ; Search time 13.8875 Seconds
(without alignments)
949.162 Million cell updates/sec

Title: US-09-674-379a-13
Perfect score: 2533
Sequence: 1 MPGRIRITVITLALCLPSP.....INFRGSSVIRLRIVSYQYF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2533	100.0	448	2	US-08-884-072-1
2	2533	100.0	448	4	US-09-212-168-1
3	1791	70.7	337	4	US-09-188-930-186
4	1289	50.9	443	2	US-08-833-963C-2
5	1289	50.9	443	3	US-08-980-514-1
6	963.5	38.0	387	2	US-08-884-072-5
7	963.5	38.0	387	2	US-08-833-963C-9
8	963.5	38.0	387	3	US-08-980-514-3
9	963.5	38.0	387	4	US-09-212-168-5
10	912.5	36.0	274	4	US-09-188-930-336
11	544	21.5	1394	6	5177197-30
12	475	18.8	638	2	US-08-897-443-1
13	463	18.3	1833	3	US-08-479-722B-2
14	463	18.3	1833	5	PCR-US95-02251-18
15	461	18.2	1251	5	PCR-US95-02251-3
16	461	18.2	1252	1	US-08-199-780-3
17	461	18.2	1252	1	US-08-316-650-3
18	461	18.2	1253	2	US-08-479-722B-4
19	452	17.8	956	2	US-08-897-443-3
20	392	15.5	886	3	US-09-110-116-3
21	377	14.9	2471	1	US-08-185-432-16
22	377	14.9	2471	1	US-08-083-590A-19
23	377	14.9	2471	3	US-08-532-384-19
24	377	14.9	2471	4	US-08-899-232-1
25	375	14.8	810	2	US-08-820-170A-34
26	375	14.8	810	3	US-09-055-699-34
27	375	14.8	810	4	US-09-273-565-34

28	375	14.8	810	4	US-09-565-538-34	Sequence 34, Appl
29	375	14.8	810	4	US-09-661-468-34	Sequence 34, Appl
30	366	14.4	816	2	US-08-820-170A-37	Sequence 37, Appl
31	366	14.4	816	3	US-09-055-699-37	Sequence 37, Appl
32	366	14.4	816	4	US-09-273-565-37	Sequence 37, Appl
33	366	14.4	816	4	US-09-565-538-37	Sequence 37, Appl
34	366	14.4	816	4	US-09-661-468-37	Sequence 37, Appl
35	360.5	14.2	1964	4	US-09-467-997-1	Sequence 1, Appl
36	359	14.2	2703	1	US-08-185-432-19	Sequence 19, Appl
37	359	14.2	2703	4	US-08-899-232-4	Sequence 4, Appl
38	350.5	13.8	652	2	US-08-751-305-2	Sequence 2, Appl
39	347.5	13.7	2556	1	US-08-083-590A-20	Sequence 20, Appl
40	347.5	13.7	2556	3	US-08-532-384-20	Sequence 20, Appl
41	346	13.7	2523	1	US-08-185-432-18	Sequence 18, Appl
42	346	13.7	2523	4	US-08-899-232-3	Sequence 3, Appl
43	344.5	13.6	2556	1	US-08-185-432-17	Sequence 17, Appl
44	344.5	13.6	2556	4	US-08-899-232-2	Sequence 2, Appl
45	340.5	13.4	673	1	US-08-282-141-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-1
; Sequence 1, Application US/08884072
; Patent No. 5872234
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,072
; FILING DATE: Herewith
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 448 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNUTO1
; CLONE: 45517
; US-08-884-072-1
Query Match 100.0%; Score 2533; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 2.6e-193;

Db 785 PLECNPGYTLGSDKCDVIDECCOKONGCS--HRCSTNGSGFKC----SCPPGYELSDQ 838
 QY 331 R-CM---CPAENPGC 341
 Db 839 KTCODIDECDODKTSK 854

RESULT 15

096SC3 PRELIMINARY; PRT: 2673 AA.
 AC 096SC3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Fibulin-6 (Fragment).
 GN Fibulin-6.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBL_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MELANOMA;
 RA Kostka G., Timpl R.;
 RT "Partial sequence of fibulin-6 with a c-terminal region related to
 domain II and III of the fibulin family."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ306906; CAC37630.1; -;
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000875; Cectropin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00008; EGF_5.
 DR Pfam: PF00047; Ig_17.
 DR Pfam: PF00090; TSP_1; 6.
 DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_5.
 DR PROSITE: PS00268; CECROPIN; UNKNOWN_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
 DR PROSITE: PS01187; EGF_CA; UNKNOWN_8.
 DR PROSITE: PS50092; TSP1; 6.
 DR NON_TER 1
 FT SEQUENCE 2673 AA; 291017 MW; BEAC30B8340E272 CRC64;

Query Match 21.68; Score 548; DB 4; Length 2673;
 Best Local Similarity 26.98; Pred. No. 7.4e-46;

Matches 144; Conservative 66; Mismatches 203; Indels 122; Gaps 17;

QY 16 CLPSPGNAQACTNGFPLDRSQGLDIDECRTIPEACRGDMGVNONGYLCTPRTNPV 75
 Db 2159 CHNMGTYTSCPKGLTIADGRICODIDECAALGRHCHAGODCDNTIGSTRVRCGSG 2218
 QY 76 YRGPTSNPYSTPYSGPYPAAPLADPNPTISRPILICRFGYQMDENQCVADVDECATDS 135
 Db 2219 FRRT-SDGLSCODINEQESSPCHQRCFNAIGSFHGCCEPGYQL-KGRKCDVNECR--Q 2274
 QY 136 HQCNPOICINTGGGYTC--SCIDGYWLE-GCCLDIDECRYGQ-----QLCANVPGSY 188
 Db 2275 NVCRPDCHKNTGRGYCIDLCPNGMTKAENGTCIDIDECRGTQHCRCRYNQCENRTGSY 2334
 QY 189 SCTCNGFTLNEGRSCDVNEC-ATENPCVQTCVNTYGSFICRCDDPGYELEEDVHCS- 246
 Db 2335 RCYCPRGYSQSGVGRPCMDINECRQVPRCAHQCSNTPGSFKCICPPGHLIGDCKSCAG 2394
 QY 247 ----- 246
 Db 2395 LERLPVNGTQSYNLA RFSPVNNYQPOQH YRQYSHLYSSYSEYRNSRTSLSTRRTIR 2454
 QY 247 -----DMDECSFSEFLCOHCECVNQPRTYFCSCPPGYILLDDNRSCQDINE 291
 Db 2455 KTCPEGSEASHDPCVDIDECENTD-ACQHECKNTFGSYOCICPPGYQLTHNGKTCODIDE 2513

QY 292 CEHRNHTCMLOQTCYNLGGFKCIDPICEPEYLRIIS-DNRKM--CPAENPGCDQPFETI 348
 Db 2514 CLEQNVHCGPNRMCFFNMRSGYQCID-TPCPPNVRORDVSGFCCLKNCPPNDECALSPYAL 2572
 QY 349 LYRDMDVSGRYPADIFQMOATR---YGAAYI-----FOIKSGNEGREFYMR 395
 Db 2573 EYKLVSIPGIAINODLIRLVAAYTODGVHMRPTFTFLMWDEQTVPFALRDEN----- 2624
 QY 396 QTGPISATLVMTRPIKPR--ETOLDLMTVNTVINFRGSSVRLRIYSOYPE 448
 Db 2625 ----LKGVYVYTRPLLEAEYTRMRVRASSYSANGTIERQYTFIV--YIANSAPY 2673

Search completed: July 3, 2003, 18:21:57
 Job time : 35.9041 secs

```

QY 69 --IP--RTNPYRGP---YSNPYSPPGPPAAPP-----98
DB 227 RLVPHRHRANRGNARRRRRADDYSR--AGEYRASQANTFEGCPMGHLPQHGHCVDIDE 284
QY 99 -----LSAPNYPTRSRPLICRFQYOMD--ESNOCV-----126
DB 285 CATLMDDCLESORCLNTPGSFKCIKRLTSCGTGYAMDESETERNNCFLLIINTFNCKYFEV 344
QY 127 -DVDECATSHOHCNPQICINTEGYTCS-----154
DB 345 EVDVDECNLASHDQGLPYQCRNTGYSKCDAAKKGCGDGLQNMTEGYIDEVYTHNGCAGE 404
QY 155 -----CTDGYWL--LEGOCLDIDEGRYGCQOL-CANPVSYSCTGNP 194
DB 405 ECVNTPGSPRCQOKMLCAHGEVNGATGFCEDVNECQGVCSMECINLPQYTKCKRGGP 464
QY 195 GFTLNE-----DGRSCQDYNE 210
DB 465 GYEFNDAKRCEVDDECIRFAGHYCDLTAECINTIGSEFCKCKPQFOLASDGRCEDYNE 524
QY 211 CATE-NPCVQTCVNTGSGFICRCDPGELEBEDGVHCSDMDECSF-----SEFLQHECVNQ 265
DB 535 CTTGTAACQKQKCNIGSGYQICIDRGFALDPDGKCEDIDECSTIWAGSGNDLGMGCGINT 584
QY 266 PGTYFCSPGYYILDDNRSCODINECEHNNHCNLAQTCYNLGGFCIDPICEEPYL 325
DB 585 KSYLQCCPFGYKIQPDGRGVYDECA-NGECAGSKVCYNTIGSFRC-HSIDCPYTI 642
QY 326 RISDN-----CMCPAENGCGRDQPTILYRDMVYSGRSVP-----362
DB 643 HSLNKNRCNRQPSACGLPEE--CSKVPLFTFYQFISL-ARAVPISHRPAITLFKVS 697
QY 363 ----ADI--FQMAQTRYPGAYYIPIQKSGNEGRYMTGP--ISATLVMPRIQGP 413
DB 698 APNHADEYVNFELQKTTIYGADNVLPALIRAN-----FLQKGRKNSAVVTLRSLDGP 752
QY 414 REIOLDL 420
DB 753 QTVKIQDL 759

```

RESULT 14

QYVS89 PRELIMINARY; PRT; 1394 AA.

AC QYVS89: 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)

DT CG7526 protein (Fragment).

OS Drosophila melanogaster (Fruit fly)

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydriidae; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

RA BRUNSON R.C., ROGERS Y.H.C., BLAZER R.G., CHAMPE M., FLEIFER B.D.,

RA WAN K.H., DOYLE C., BAXTER E.G., HELT G., NELSON C.R., MIKLOS G.L.G.,

RA ABRIEL J.F., ABBAYANI A., AN H.-U., ANDREWS-PINAKOCH C., BALDWIN D.,

RA BEESON K.Y., BENOS P.V., BERMAN B.P., BHANDARI D., BOLSHAKOV S.,

RA BORISOVA D., BOTCHAN M.R., BOUCK J., BROKSTEIN P., BROTLIER P.,

RA BURTLIS K.C., BUSAN D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,

RA CHERRY J.M., CAWLEY S., DAHLKE C., DAVENPORT L.B., DAVIES P.,

RA DE PABLOS B., DELCHER A., DENG Z., MAYS A.D., DEW I., DIETZ S.M.,

```

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jamal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J.J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.A., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RA Science 287:2185-2195(2000).
RA EMBL: AE003558; AAF0538.1; -.
RA HSSP: P00736; IAP0.
RA FlyBase: FBgn0035798; CG7526.
RA InterPro: IPR000152; ASX_Hydroxyl.
RA InterPro: IPR000561; EGF_Like.
RA InterPro: IPR001881; EGF_Ca.
RA InterPro: IPR003410; Hyalin.
RA InterPro: IPR000436; Sush1_SCR_CCP.
RA InterPro: IPR001491; Thrombomodulin.
RA Pfam: PF00008; EGF_11.
RA Pfam: PF00084; Sush1_2.
RA PRINTS: PR00907; THROMBOMODULN.
RA SMART: SM00032; CCP_2.
RA SMART: SM00179; EGF_Ca_9.
RA SMART: SM00001; EGF_Like_5.
RA PROSITE: PS00010; ASX_HYDROXYL_8.
RA PROSITE: PS01186; EGF_2_9.
RA PROSITE: PS01187; EGF_CA_10.
RA Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
RA NON_TER 1394 1394
RA FT SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;

```

Query Match 22.0%; Score 558.5; DB 5; Length 1394;

Best local Similarity 32.4%; Pred. No. 3e-47;

Matches 122; Conservative 42; Mismatches 135; Indels 77; Gaps 13;

```

QY 15 LCIPSPGNMAOCTNGD---LDROSGCLDIDECRTITPEACRGMMKCVNONGYLCTPR 71
DB 507 VCNRLPSYSGICAGAGELLKLDIRGCFDIDECSTQTHGCSQDLNENLSTYCL-- 564
QY 72 TNVPYRGPYNPSTYSPGPPAAPPPLASPNYPTISRPLICRFQYOMDESNOQVDVDC 131
DB 565 CPFGYALGLDNHIVTSLNSFTIDTSTSEPTS-----ATTCIDIDEC 606
QY 132 ATDSDHCNPQICINTEGYTCSCTDGYWLEGO--CIDIDEC--RYGYCOOLCANVPS 187
DB 607 SLANGNC--SHFQCNREGGFGQACPLGVALSEDMRTCODIDECIDLSNQCQSLCINQPG 664
QY 188 YSTCTNPGFTLNDGRSCQVNECATE-NPCVQCVNNTGSGFICRCDPGELEBEDGVHKS 246
DB 665 FACACETGFEFTDGFECACIDECSDYGNCSDCITNLGTHACCEGYELAKKLSCL 724
QY 247 DMEDEC-SFSEFLQHECVNPGTYFCSPGYIILDDNRSCQ-----287
DB 725 DVDECAGLSLSGSGSHEDINNAGTFEGCCPLGIYILNDDGRSSSPALVGCPGTRASADGA 784
QY 288 -----DINECEHRNHTCNLAQTCYNLGGFCIDPICEEPYLRISDN 330

```

```

Db 387 GNDLCMGCGINTGSGYLCCGCPGKIPDERTCYDVDECA-MGECAGSDKCVNTLGSFK 445
OY 314 CIDPICEEPLYLISDR-----C-MCPAENPC-RDOPETILYRDMDVYSGRSV- 361
Db 446 C-HSIDCPNYIHDLSLNKNOIADGYSICIKVCTEDTCLGNHTREVIYQFAVPSLKII 504
OY 362 -PADI-----FOMATTRYPGAYYIFQIKSGNEGREFYKRCQGPICATLVMRPI 410
Db 505 SPIEVSRIYTHMGVPSVVDYNDLYGGRHFRIVOERNIG-----IVQLVKPI 551
OY 411 KGPREFOLDLEMTVTNVTINFRGSSVIR-----LRIVSOYF 448
Db 552 SGP-----TVEIKVINHTKSRFGVILAFNEAIIISVSTPF 589

RESULT 12
OY 090H16 PRELIMINARY; PRT; 554 AA.
ID 090H16;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DJ162H14.1 (Fibulin 1) (Fragment).
GN FBLN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68047; CAB62995.1; -.
DR HSSP; P35555; IEMN.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 3.
DR Pfam; PF00008; EGF; 6.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_Ca; 4.
DR SMART; SM00001; EGF-like; 5.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 3.
DR EGF-like domain; Glycoprotein; Hydroxylation.
FT NON TER
SQ SEQUENCE 554 AA; 59767 MW; FE285184599A2982 CRC64;

Query Match 23.0%; Score 583; DB 4; Length 554;
Best Local Similarity 36.1%; Pred. No. 3.4e-50;
Matches 116; Conservative 47; Mismatches 102; Indels 56; Gaps 13;

OY 16 CLSPGNAOACTNGFDLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCPRTNVP 75
Db 261 CRP-----KLCKCKGFLTD-ALGNICIGINECLISAPCPITGHTCINTGSSYTC----- 307
OY 76 YRCPYNSPYSPYSGPYPAAPRISAPYPIISRILTRFGQM-DESNQCVDVDECATD 134
Db 308 -----QKNVPR-----CGRGHINEGTRCVGVDECAAP 336
OY 135 SHOCNPTQICINTEGGYCTSCDGYWL-LEGOCLDIDEC-RY-GYCOOLCANVPGS 189
Db 337 AEPCKGKRCVNSPGSRCECKTGYFDGISRMCVGVNECQRYERLCHKCENTLGLSTL 396
OY 190 CTNCPGFTLLEDGRSCDVNECATENPCVOTCVNTYSGFICRDPGYLEE-DGVHCS 248
Db 397 CQSVGFRLSVDRSCGINECS-SPQSECANVYSGCYCRGQYLSVDVGYTCGSI 455
OY 249 DECSF--SEPLCQHECVNPGTYFCSCP-PGYILLDNRSODINECHRNHTCNLQOTC 305

```

```

Db 456 DECAFLTGTGICSYRCINIPGSCSPSSGYRLAPNGRCOGIDEVCTIHNCSINETC 515
OY 306 YNLGGFCRCLDIPRCEPYLR 326
Db 516 FNIQGFRL-APECPENTRR 535

RESULT 13
OY 018026 PRELIMINARY; PRT; 798 AA.
ID 018026;
AC 018026; Q20903;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F56H11.1 protein.
GN F56H11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitodea;
OC Rhabdilitae; Pelodetrinae; Caenorhabdilitis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Percy C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaughan M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sprout J., Wohldman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.*
RT Nature 368:32-38(1994).
RN [3]
RP SEQUENCE FROM N.A.
RA Lloyd C.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68749; CA92962.1; -.
DR EMBL; Z68219; CA92962.1; JOINED.
DR EMBL; Z68219; CA92483.1; -.
DR EMBL; Z68749; CA92483.1; JOINED.
DR HSSP; P16109; IFSB.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF01821; ANATO; 2.
DR Pfam; PF00008; EGF; 5.
DR SMART; SM00104; ANATO; 2.
DR SMART; SM00179; EGF_Ca; 4.
DR SMART; SM00001; EGF-like; 6.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_Ca; 8.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 798 AA; 87205 MW; 3BF1EE9BD54D8BF9 CRC64;

Query Match 22.8%; Score 577.5; DB 5; Length 798;
Best Local Similarity 26.7%; Pred. No. 1.9e-49;
Matches 162; Conservative 50; Mismatches 180; Indels 215; Gaps 23;

OY 15 LCLSPG-NAOACTNGFDLDROSGOCLDIDECRTIPEACRGDMCVNONGYLCC----- 68
Db 167 LCHDRGERVEVSCSGFDLAPDGMACVDRNECLTRQSPCTQSDCVNTIGTYICQRRIS 226

```


Db 539 FNCTENIOGFCRLS-LECPENRYKSGDVRLEKTDITRICKSCRPNDVNOVLDPVHTI 597
 QY 349 LYRDMVYSGRSV--PADIFOMQA-TTRYRG--AYVFOITSGNEGREF-----YMRGTG 398
 Db 598 SRVIVSLPTRETTREPELITFLRAITPTYPANQADILFDITEGMLRESFDIIRKM--DG 655
 QY 399 PISATLVMTPIKPIKPREIQLDEM-ITVNTVYINFRGSSVIRLRIVYSQPF 448
 Db 656 MRYGVVROYRPIVGFPHALLKLENNYMGVSHR--NIYVNHIFVSEYWF 704

RESULT 8

Q913V7 PRELIMINARY; PRT; 576 AA.
 AC Q913V7;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Hypothetical 63.3 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Mambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL050095; CAB43267.1; -
 DR HSSP: F00736; IAPQ.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00179; EGF_CA: 8.
 DR SMART: SM00001; EGF_Like: 2.
 DR PROSITE: PS00010; ASX-HYDROXYL; 4.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA: 9.
 DR Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation;
 KW Hypothetical protein; Repeat.
 FT NON_TER 1
 SO SEQUENCE 576 AA; 63274 MW; FF2F0E9B185D8AC3 CRC64;

Query Match 28.1%; Score 711.5; DB 4; Length 576;
 Best Local Similarity 34.3%; Pred. No. 3.9e-63;
 Matches 146; Conservative 61; Mismatches 156; Indels 63; Gaps 12;

QY 15 LCLSPGN-----AQACGTNGFDLDROSGGCLDIDECTRIPACRGDMGVNONGYLCTIP 70
 Db 172 LCONTKSGFYCOARORCMDF-LDPEGNCVDINECTSLSEPCRPFGSCINTVGSYTC-- 228
 QY 71 RTNPNYRQPNYSTPYSGPYPAAPPLSAPNYPTISRLICRGFYQ--MDESNQCVNDV 129
 Db 229 -----QKNVPLICARVYHSDGTCVNDV 252
 QY 130 ECATDSHQCNPTQICINTGGYTCSTDGWY--LLEGCLDIDECCRYG--YCOQLCANV 184
 Db 253 ECEFGVHRSGEGYCHNLPGSYRCCKAGFORDAFGRCIDVNECHASGRCLQHCNCEHT 312
 QY 185 PGYSICNPGFTLNEDGRSCODVNECATENPCVQTCVNTYFSFICRCDPGYLEEDGVH 244
 Db 313 LGSYRCASGFLLAADKRCEDVNECEAQR-CSQECANITYSGCYCRQYQALADGHR 371
 QY 245 CSDMDECS-FSEFICQHECVNQPFTYSCSP-PGYILLDDNRSQDINCEHRNHTCNQ 302
 Db 372 CTDLDECAQAGAILCTPRCLVNPSSYQACPEQGYTPMRANGSKCDVDCALGTHNCSBA 431
 QY 303 QTCVNLGGFKCIPDIRCEEPYLRISDNRCMCAPENPCRD-----QPTILLYRDMVY 356
 Db 432 ETCNIGGSRCL-RFECPPNYVGVSKTKC-----ERTCHDLFECQNSPARITHQNLQ 486

QY 357 SGRSVADPFDIOMQATTRYPGAYYIFQIKSGNEGHEFYMRGTGPISATLVMTPIKGPRI 416
 Db 487 TGLVPAHIFRLGPAPAFGTDTALNTIKGNEBGFETRLRNATGVVYIQRAVLEPRDF 546
 QY 417 OLDLEM 422
 Db 547 ALDVEH 552

RESULT 9

Q9HB05 PRELIMINARY; PRT; 495 AA.
 AC Q9HB05;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Hypothetical 54.3 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxId-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gu J.R., Man D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
 RA Qiu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
 RA Yu J., Han L.H.;
 RT "Novel Human cDNA clones with function of inhibiting cancer cell
 growth."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217999; AAG17241.1; -
 DR HSSP: P35555; IEMN.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR SMART: SM00181; EGF_9.
 DR SMART: SM00179; EGF_CA: 9.
 DR SMART: SM00001; EGF_Like: 1.
 DR PROSITE: PS00010; ASX-HYDROXYL; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA: 8.
 KW EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
 SO SEQUENCE 495 AA; 54340 MW; C40434EB6C82E3D70 CRC64;

Query Match 26.7%; Score 676; DB 4; Length 495;
 Best Local Similarity 35.8%; Pred. No. 1.3e-59;
 Matches 139; Conservative 59; Mismatches 128; Indels 62; Gaps 16;

QY 16 CLPSPGNAOACTNGFDLDROSGGCLDIDECTRIPACRGDMGVNONGYLCTIPNTVY 75
 Db 158 CRP-----KLOCKSGFTD-ALGNCLDINECLISAPCPIGHTCINTBSYTC----- 204
 QY 76 YRGPSNPNYSTPYSGPYPAAPPLSAPNYPTISRLICRGFYQ--MDESNQCVNDV 134
 Db 205 -----QKNVPLN-----CGRGVHNEBGRVNDVDECAFP 233
 QY 135 SHQCNPTQICINTGGYTCSTDGWY--LLEGCLDIDECCRYG--YCOQLCANVPGSY 189
 Db 234 AEPGCKHRCVNSGSEFCECKTGTFDGIISRMCDVNECEQYRGLRGHKCENTGYSYL 293
 QY 190 CTCPNGFTLNEDGRSCODVNECATENPCVQTCVNTYFSFICRCDPGYLEEDGVH 248
 Db 294 CSCGVRLSTVDSGSCDINECCS-SFQSCQCANVYSGCYCRGQYQLSDVDGYTCEDI 352
 QY 249 DECSF--SEPLCQHECVNQPFTYSCSP-PGYILLDDNRSQDINECHRNHTCNQ 305
 Db 353 DECAPLPTGHCISYRCINIPSGFQSCPSGSRILAPNRNQCDDIDECVYTGINCINERG 412
 QY 306 YNLGGFKCIPDIRCEEPYLRISDNRC-MCPA-ENPGRDQPTLLYRDMVYSGRSVPA 363
 Db 413 FNIQGGRCCL-AECPENYRSATRCERLCHENRECSKLPRLITYVHLSPFTIOAPA 471
 QY 364 DIFOMQATTRYPGAYYIFQIKSGNEGHE 391

```

QY 364 DIFOMQATRRPGAYYIFQIKSGNEGREFYMRQGPISATLWMTRPKPREIOLDLMI 423
DB 602 VIFRMCPSAVNGDSNQALITGNEBGFITTKKVSFHSVVALTKPVPEDL-----LT 656
QY 424 TVNTVINFRG---SSVIRLRIYVS 444
DB 657 TVKMDLSRHGTVSSFFAKLFIFVS 680

RESULT 6
QY 099K58 PRELIMINARY; PRT; 1174 AA.
AC 099K58;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to fibulin 2.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strusberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005443; ANH05443.1; -.
DR HSSP: P00736; IAPQ.
DR MGD: MGI:95488; EDln2.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF01821; ANATO; 2.
DR Pfam: PF00008; EGF; 6.
DR SMART: SM00104; ANATO; 3.
DR SMART: SM00179; EGF_CA; 9.
DR SMART: SM00181; EGF; 11.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.
DR PROSITE: PS01186; EGF_2; 5.
DR PROSITE: PS01187; EGF_CA; 9.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 1174 AA; 126460 MW; 8D628AC710FBA6B8 CRC64;
SQ

Query Match 29.0%; Score 735.5; DB 11; Length 1174;
Best Local Similarity 35.3%; Pred. No. 3.4e-65;
Matches 146; Conservative 61; Mismatches 148; Indels 59; Gaps 11;

```

```

DB 1042 L-RFDCPPNVYRVSETKC-----ERTTCODITECQTSPARTHYOLNFOGLLPAHIFRI 1096
QY 369 QATRRPGAYYIFQIKSGNEGREFYMRQGPISATLWMTRPKPREIOLDLMI 422
DB 1097 GPAPAFAGDTISLITTKGNEGYFVTRNLNAYTVGSLSORSVLEPRDALDEM 1150

RESULT 7
QY 073774 PRELIMINARY; PRT; 704 AA.
AC 073774;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Fibulin-1, isoform D.
GN FBLN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99120531; PubMed-9923656;
RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
RT Identification of chicken and C. elegans fibulin-1 homologs and
RT characterization of the C. elegans fibulin-1 gene.;
RL Matrix Biol. 17:635-646(1998).
DR EMBL: AF051399; AAC05387.1; -.
DR HSSP: P00742; IHCG.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF01821; ANATO; 2.
DR Pfam: PF00008; EGF; 6.
DR SMART: SM00104; ANATO; 3.
DR SMART: SM00179; EGF_CA; 8.
DR SMART: SM00001; EGF_Like; 1.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 8.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 704 AA; 76137 MW; D47D5A3D05E92932 CRC64;
SQ

Query Match 28.5%; Score 722; DB 13; Length 704;
Best Local Similarity 34.6%; Pred. No. 4.4e-64;
Matches 163; Conservative 68; Mismatches 158; Indels 82; Gaps 22;

```



```

OY 56 DMWCVNONGYLCTIPRTNPNYRGPNYSPTSPYSGPYPAAPPLISAPNPTISRLICRF 115
DB 68 EMKCIHHNGGYLCLPNSAAVINDLHG-----EGP-PPVPPAQHNP-----PCPP 111
OY 116 GYOMDESNOCVDDECATOSHOCNPTQICINTEGGYCSCTDGYWLEGGCLDIDECRYG 175
DB 112 GYEPDQDSCVDDECAQALHDDRPSDDCHNLSSGYCTCCTCPDGRKGPEECVDIDECRYR 171
OY 176 YCOQLCANPGYSCITCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRCDPG 235
DB 172 YCOHRCVNLPGSFRCCCEPFGOLGPNRSCVDVNECMGAPCEQRCSNXYCTFLCRHQG 231
OY 236 YLEEDGVGSCDMDDECSFSEFLCOHCVNPGTYFCGCPGTYILLDNRSQODINEGHR 295
DB 232 YELHNDGFCSDIDECISSTYLCOYKCVNEGRFSCHCPGYQLL-ATRLCQDIDECESG 290
OY 296 NHTCNLOOTCYNLOGGFKCIDPICEPEYLRISDNRCMPAENPGCRDOPFTILYRMDV 355
DB 291 AHOCSEAOICVNFHGGYRCVDTNRCVEPTQVSENLCLCPASNPLCREQSSIVHRYMTI 350
OY 356 VSGRVPADIFOMQATRRPGAYITQISGNGREYMRQTFISATLVTRPIKGPPE 415
DB 351 TSERRRPADVFOIQATSYVPGAYNAFOIRAGNSQGFYIRQINNVSAMLVLRDPVGPPE 410
OY 416 IQLDLEMITVNTVINRGSSVIRLRIVYSQYP 448
DB 411 YVLDEMTVNTMSLMSSYRASSVLRITVFGAYTF 443

```

RESULT 4
0922K8 PRELIMINARY: PRT: 685 AA.

```

ID 0922K8 AC 0922K8 DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similiar to fibulin 1.
GN FBLN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
BL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DB EMBL: BC007140; AA07140.1; -.
DB MGI: 95487; Fbln1.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx-hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR Pfam: PF01821; ANATO; 3.
DR Pfam: PF00008; EGF; 6.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN_3.
DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE: PS00010; ASX-HYDROXYL; UNKNOWN_4.
DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
DR PROSITE: PS01187; EGF-Ca; UNKNOWN_8.
SQ SEQUENCE 685 AA; 75283 MW; EFOD77D7F66B73B8 CRC64;

```

Query Match 29.3%; Score 743; DB 11; Length 685;
Best Local Similarity 36.1%; Pred. No. 3.2e-66;
Matches 160; Conservative 65; Mismatches 154; Indels 64; Gaps 17;

```

OY 16 CLSPGNAQAQCTNGFDLDROSQCCLDIDECRTIEACRGDMCVNONGYLCTIPRTNPV 75
DB 290 CRP-----KLCCKSGFIOD-ALGNCIDINECLISAPCPVGOTCINTBESYTC----- 336
OY 76 YRGYSNPNYSPYSGPYPAAPPLISAPNPTISRLICRFQYOM-DESNQCVNDECATID 134
DB 337 -----QKNVNP-----GGRGYHLINEGTRCVNDECSPP 365

```

```

OY 135 SHOCNPTQICINTEGGYCSCTDGYWL--LEGQCCLDIDEC-RY--GYCOQLCANPGSYS 189
DB 366 AEPCKGHHCLNLSPGSFRECKAGFYFDGJSRRCVNDNEQGRGLRGKHCENTPSSFH 425
OY 190 CTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRCDPGYLEE-DGVHCSDM 248
DB 426 CCSAGFRLSVGRSCEDVNEC-LNSPCSQECANVGSYQCYCRGYQLSDVDGVCEDI 484
OY 249 DECSF--SEFLCOHCVNPGTYFCSCP-PGYILLDNRSQODINEGHNHNCNLOOTC 305
DB 485 DECAPLTGHTICSYKINIPGSFQCSPPSGYRLAPGRNCQDIDECVTGTHNCISNETC 544
OY 306 YNLOGGFKCIDPICEPEYLRISDNRC-MCPA-ENPGCRDOPFTILYRMDVVGSRVPA 363
DB 545 FNIGSFRLS-FECPENYRRSADTCERLPCHENEGCPRLPLRITYHLSFTNIOVPA 603
OY 364 DIFOMQATRRYGAYITQISGNGREYMRQTFISATLVTRPIKGPPEIQLDLEW- 422
DB 604 VFERMGPSAVPGDSMQALITAGNEGFFTRKRVSHSGVALLTKPIPEPRDILLVKMD 663
OY 423 -ITVNTVINRGSSVIRLRIVYS 444
DB 664 LYRHGTVSF-----VAKLFIYVS 682

```

RESULT 5
08TBH8 PRELIMINARY: PRT: 683 AA.

```

ID 08TBH8 AC 08TBH8 DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Fibulin 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-BRAIN;
RA Strausberg R.;
BL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DB EMBL: BC022497; AA022497.1; -.
SQ SEQUENCE 683 AA; 74423 MW; 2665A3961B6403B4 CRC64;

```

Query Match 29.1%; Score 737; DB 4; Length 683;
Best Local Similarity 35.4%; Pred. No. 1.3e-65;
Matches 157; Conservative 67; Mismatches 154; Indels 66; Gaps 17;

```

OY 16 CLSPGNAQAQCTNGFDLDROSQCCLDIDECRTIEACRGDMCVNONGYLCTIPRTNPV 75
DB 288 CRP-----KLCCKSGFIOD-ALGNCIDINECLISAPCPIGHTCINTBESYTC----- 334
OY 76 YRGYSNPNYSPYSGPYPAAPPLISAPNPTISRLICRFQYOM-DESNQCVNDECATID 134
DB 335 -----QKNVNP-----GGRGYHLINEGTRCVNDECAPP 363
OY 135 SHOCNPTQICINTEGGYCSCTDGYWL--LEGQCCLDIDEC-RY--GYCOQLCANVGSYS 189
DB 364 AEPCKGHHCLNLSPGSFRECKAGFYFDGJSRRCVNDNEQGRGLRGKHCENTPSSFH 425
OY 190 CTCNPGFTLNEDGRSCODVNECATENPCVOTCVNTYGSFTICRCDPGYLEE-DGVHCSDM 248
DB 424 CCSVGFRSLVDGSCEDINECSS-SPCSQECANVGSYQCYCRGYQLSDVDGVCEDI 482
OY 249 DECSF--SEFLCOHCVNPGTYFCSCP-PGYILLDNRSQODINEGHNHNCNLOOTC 305
DB 483 DECAPLTGHTICSYKINIPGSFQCSPPSGYRLAPNNSCQDIDECVTGTHNCISNETC 542
OY 306 YNLOGGFKCIDPICEPEYLRISDNRC-MCPA-ENPGCRDOPFTILYRMDVVGSRVPA 363
DB 543 FNIGGGRCL-AEPCEPNYRSATRCERLPCHENRECSKLPRLITYHLSFTNIOAPA 601

```

```

QY 56 DMMCVNONGGLTCLPRTNPVYRGSPNPSYSPYSGYPAAPALSPNPTISRPLICRF 115
    : : | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 68 EMKCTNNHGGTCLDRLRSAAVINDLHG-----EGP-PPVPVPAQHNP-----PCPP 111
QY 116 GYQMDNESNOCVADVDECATDSDHOCNPTQITCNTEGGTGCSTGUYMLLEGOCIDIDECRYG 175
    | | : : | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 112 GYEPDPODSCVDYVDECAALHCCRISQDCHNLPGSYQCTCPDGRYKRIIGPCVDIDECRYR 171
QY 176 YCQOLCANPSSYSCTCNBPFTLNEDGRSCQDVNECATENPCVQTCVNTYGSFICRDPG 235
    | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 172 YCQHRCVNLPGSFRQCEPFGFOLGPNRRSCVDVNECDMACPCBQRCFNSYGTFLCRHOG 231
QY 236 YELEBDGYHNGDMDSCSFSEFCQHECVNQDPETPCSCPRGYLLDNDNSCODINCEHR 255
    | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 232 YELHRDGSQSDIDECSSYSYLQYKCVNEPFRFSCHQGYOLL-ATRLCQDIDEECSG 290
QY 296 NHTCNLAOTCYNLQGFKCIDPICEEPLRLISDNCMCPAENPGCRDQPFLLIYRDMV 355
    : : | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 291 AHQSEAOATCVNHFGRYRCVDINRCVEPIQYSENNCKCPASNPCLREQPSYVHKRMTI 350
QY 356 VSGRSVPADITFOMQATTRYPCGAYVYIFQIKSGNEGREFYRKQGTGPISATVLMTRPIKPRE 415
    | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 351 TSERSVPADVFOIQATSYVPCGAYNAFOIIRAGNSOGFYTROIINNVASAMVLARPYGPRE 410
QY 416 IQOLDLEITVTVINFRGSSVIRLRIYQYRF 448
    | | | | | | | | : : | | | | | | | | : : | | | | | | | |
Db 411 YVLDELEVMTNLSMYRASSVLRVLVEGAYLRF 443

```

RESULT 2
Q9JM06
ID Q9JM06 PRELIMINARY; PRT; 443 AA

01-OCT-2000 (TREMblrel. 15, Created)
 01-OCT-2000 (TREMblrel. 15, last sequence update)
 01-JUN-2002 (TREMblrel. 21, last annotation update)
 EGF-containing fibulin-like extracellular matrix protein 2.
 EFEMP2.
 OS Mus musculus (Mouse) .
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 []
 SEQUENCE FROM N.A.
 RP MEDLINE=20435063; PubMed=10962184;
 RX Katsanis N., Venable S., Smith J.R., Lupski J.R.;
 RA "Isolation of a paralog of the Doyne honeycomb retinal dystrophy gene
 from the multiple retinopathy critical region on 11q13."
 Hum. Genet. 106:66-72(2000)
 EMBL: AF109122; AAF65189.1; -.
 DR HSSP: P00736; IAPQ.
 DR MGD: MGI:1891209; Efemp2.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF Ca.
 DR InterPro: IPR001491; ThnbdomduLn.
 Pfam: PF00008; EGF_4.
 DR PRINTS: PR00907; THNBDOMDU LN.
 DR SMART: SM00179; EGF_CA_4.
 DR SMART: SM00001; EGF-like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS0186; EGF_2; 4.
 DR PROSITE: PS0187; EGF_CA_6.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
 KW Matrix protein; Repeat.
 QM SEQUENCE 443 AA; 49452 MW; 5AEC2A91048B336A CRC64;

Query Match	50.7%	Score 1283	DB 11	Length 443
Best Local Similarity	49.9%	Pred. No. 6.6e-121		
Matches	226	Conservative	70	Mismatches 135; Indels 22; Gaps 4
by	1	MPGIKRLITVTLALCPSPGNAO-----A	ACTGNGFDLDRSGGCLDIDECTTPEACNG	55

[illegible]

RESULT 3
Q9H3D5
ID Q9H3D5 PRELIMINARY; PRT; 443 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Fibulin-like extracellular matrix protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi,
 CC Mammalia; Euteria; Primates; Catarrhini; Hominae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RA Seibold S., Marx M.,
 RT "Cloning of a new fibulin-like gene."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DB EMBL; AF134486; AAC645245.1; -
 DB HSP; P35555.1EMN.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001491; Thymomodulin.
 PF Pfam; PF00008; EGF; 3.
 DR PRINTS; PR00907; THYMOMODULN.
 DR SMART; SM00181; EGF; 5.
 DR SMART; SM00179; EGF_CA; 6.
 DR SMART; SM00001; EGF_like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 5.
 KW EGF-like domain; Glycoprotein; Hydroxylation; Matrix protein.
 SQ SEQUENCE 443 Aa; 49353 MW; D91784BF36A8A060 CRC64;

	Query Match	50.0%	Score 1266;	DB 47	length 443;
	Best Local Similarity	49.28%	Pred. NO. 3,4e-119;		
	Matches	223;	Conservative	74;	Mismatches 134; Indels 22; Gaps 5
OY	1	MPGIKRIITVTTLALCPSPGNAQ-----AOCNCFDDLDROSGOCLDIDECATPEACRG	55		
db	8	LPGSLLMALLLILLLGSAPDSDEPDSYETEDDYENDMDPSOHCRDNVECTTIEACRG	67		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 15:34:32 ; Search time 32.4041 Seconds

(Without alignments)
2848.683 Million cell updates/sec

Title: US-09-674-379A-13

Perfect score: 2533

Sequence: 1 MPKIKRLVFTILALCLPSP.....INFGSSVIRLRIVSYQYF 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_rvivirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1289	50.9	443	4	096TF5
2	1283	50.7	443	11	09JMK6
3	1266	50.0	443	4	09H3D5
4	743	29.3	685	11	0922K8
5	737	29.1	683	4	08TBH8
6	735.5	29.0	1174	11	099K58
7	722	28.5	704	13	073774
8	711.5	28.1	576	4	09Y3V7
9	676	26.7	495	4	09HB05
10	631	24.9	681	13	042182
11	596	23.5	589	5	09T2S1
12	583	23.0	554	4	09UH16
13	577.5	22.8	798	5	018026
14	558.5	22.0	1394	5	09VSB9
15	548	21.6	2673	4	096SC3
16	544	21.5	1394	4	08TD95

17	537	21.2	5636	4	096RW7	096RW7 homo sapien
18	527.5	20.8	2872	11	096RW8	096RW8 ratuS norv
19	524.5	20.7	741	4	096K89	096K89 homo sapien
20	521.5	20.6	3857	11	088840	088840 mus musculu
21	519.5	20.5	2809	4	096JUP8	096JUP8 homo sapien
22	517	20.4	1713	11	088349	088349 mus musculu
23	515.5	20.4	708	13	P87363	P87363 gallus gall
24	515.5	20.4	2906	11	096WH9	096WH9 ratuS norv
25	511.5	20.2	937	5	09BLT1	09BLT1 clona Intes
26	508	20.1	188	11	08RIU8	08RIU8 mus musculu
27	495.5	19.6	1963	6	028019	028019 bos taurus
28	495	19.5	746	4	096HB9	096HB9 homo sapien
29	495	19.5	1256	4	09NS15	09NS15 homo sapien
30	495	19.5	1382	4	09H7K2	09H7K2 homo sapien
31	494.5	19.5	1095	11	060784	060784 mus musculu
32	492.5	19.4	1764	11	035806	035806 ratuS norv
33	491	19.4	517	4	09NP01	09NP01 homo sapien
34	489.5	19.3	447	5	09V4B8	09V4B8 drosophila
35	486	19.2	1511	4	075412	075412 homo sapien
36	483.5	19.1	1587	4	000508	000508 homo sapien
37	482.5	19.0	733	5	025678	025678 podocoryne
38	481.5	19.0	1062	11	060789	060789 mus musculu
39	480.5	19.0	1821	4	014767	014767 homo sapien
40	480.5	19.0	2189	5	09BI05	09BI05 elmeria ten
41	480	18.9	1289	5	08SSS3	08SSS3 dictyosteli
42	480	18.9	1664	5	09TVQ2	09TVQ2 caenorhabdi
43	475	18.8	937	4	096TF5	096TF5 homo sapien
44	468.5	18.5	669	4	075441	075441 homo sapien
45	463	18.3	1833	11	008999	008999 mus musculu

ALIGNMENTS

RESULT 1
ID 096TF5 PRELIMINARY; PRT; 443 AA.
AC 096TF5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Mutant p53 binding protein 1 (MBP1).
GN MBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka S.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tanaka S.; Sugimachi K.;
RT "Human mutant p53 binding protein (MBP1).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB030655; BAA92880.1;
DR InterPro: IPR000157; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_4.
DR PROSITE: PS01186; EGF_2; UNKNOWN_4.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_6.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_6.
SQ SEQUENCE 443 AA; 9CE175F4F388A56D CRC64;

Query Match 50.9%; Score 1289; DB 4; Length 443;

Best Local Similarity 49.9%; Pred. No. 1.6e-121;

Matches 226; Conservative 74; Mismatches 131; Indels 22; Gaps 5;
DB 8 LPSGLLMLALLLLLSASRQDSSEPDSTYECTDGYEMDPDSQHCRCVNECLTIPACKG 67

FT DOMAIN 1335 1379 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT SITE 847 849 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 304 315 BY SIMILARITY.
FT DISULFID 310 324 BY SIMILARITY.
FT DISULFID 326 339 BY SIMILARITY.
FT DISULFID 550 562 BY SIMILARITY.
FT DISULFID 557 571 BY SIMILARITY.
FT DISULFID 573 586 BY SIMILARITY.
FT DISULFID 592 604 BY SIMILARITY.
FT DISULFID 599 613 BY SIMILARITY.
FT DISULFID 615 628 BY SIMILARITY.
FT DISULFID 634 645 BY SIMILARITY.
FT DISULFID 640 654 BY SIMILARITY.
FT DISULFID 657 669 BY SIMILARITY.
FT DISULFID 675 686 BY SIMILARITY.
FT DISULFID 681 695 BY SIMILARITY.
FT DISULFID 698 709 BY SIMILARITY.
FT DISULFID 715 726 BY SIMILARITY.
FT DISULFID 721 735 BY SIMILARITY.
FT DISULFID 737 750 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 762 776 BY SIMILARITY.
FT DISULFID 778 791 BY SIMILARITY.
FT DISULFID 797 808 BY SIMILARITY.
FT DISULFID 803 817 BY SIMILARITY.
FT DISULFID 819 832 BY SIMILARITY.
FT DISULFID 838 850 BY SIMILARITY.
FT DISULFID 845 859 BY SIMILARITY.
FT DISULFID 861 873 BY SIMILARITY.
FT DISULFID 879 891 BY SIMILARITY.
FT DISULFID 885 900 BY SIMILARITY.
FT DISULFID 902 915 BY SIMILARITY.
FT DISULFID 921 933 BY SIMILARITY.
FT DISULFID 927 942 BY SIMILARITY.
FT DISULFID 944 957 BY SIMILARITY.
FT DISULFID 963 975 BY SIMILARITY.
FT DISULFID 970 984 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1109 1123 BY SIMILARITY.
FT DISULFID 1125 1138 BY SIMILARITY.
FT DISULFID 1144 1155 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1166 1179 BY SIMILARITY.
FT DISULFID 1298 1309 BY SIMILARITY.
FT DISULFID 1304 1318 BY SIMILARITY.
FT DISULFID 1320 1333 BY SIMILARITY.
FT DISULFID 1339 1354 BY SIMILARITY.
FT DISULFID 1349 1363 BY SIMILARITY.
FT DISULFID 1365 1378 BY SIMILARITY.
FT MOD_RES 647 647 HYDROXYLATION.
FT MOD_RES 810 810 HYDROXYLATION.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1394 AA; 152791 MW; DFECA81A0B2C7D1 CRC64; /FTID-CAR_000184.

Query Match 21.5%; Score 544; DB 1; Length 1394;
Best Local Similarity 29.4%; Pred. No. 7e-34;
Matches 128; Conservative 52; Mismatches 134; Indels 122; Gaps 17;

OY 16 CLPSPGNAOACTNCFDIDROSGQCIDDECTIPACRGDMACVNQNGYLCIPRTNPV 75
DB 562 CINLPVRYTCICYEGRFSEQOCVDIDECIOVHLC-SQGRCENTEGSFICI----- 614
OY 76 YRGPIYSNYSIPYSGPYPAAPPLSAPNPJTISRPLI-----CRGQXQ 118
DB 615 -----CPAGFMASEGTNCIDVDECLRPDVCGBGHCVNTVGAFRCXEYCDSGYR 662

OY 119 MDESNCQVDVDECATDSDHOCNPQICINTEGGYTC-SCTDGYWLLGQCLDIDEC-RYGY 176
DB 663 MTRGRCEIDDECLNS-TC-PDEQCVNSPGSYCVCTEGRFMNNGCLDVBELERNV 720
OY 177 CQO-ICANVPGSYSTCNPGETINEDGRSCQDVNECATENPCVO----- 219
DB 721 CANGDCSNLEGSYMCCHKGYTRTPDHRHCRDIDECQOGNLCVNGQCKNTEGSRCTGQ 780
OY 220 -----TCVMTYGSFICRDPGYLEEEDGYCSDMDECSF 253
DB 781 GYLSAKQDCEDIDECQHRHLCAHGQCRNTEGSGFCVCDGYPASGLGDHCEIDNECLE 840
OY 254 SEFLCOH-ECVNOPTGYFCSPPGYILLDDNRSCODINECHRNHTCNLAQOTCVNLAQGF 312
DB 841 DKSVCQRGDCINTAGSYDCTCPDGF-QLDDNKTCQDINECHRP-TCGPQGECLNTEGSG 898
OY 313 KCI-----DPICEEPLYRIS-----DN-----RCMC-----PAENPGCR 342
DB 899 HCVCOQGFISADGRCEIDIDECVNNTVCDSHGFCDMTAGSFRCICYGQFQAPDQGCYV 958
OY 343 DQPFITLYRDMQVYSG 358
DB 959 D-----VNECELLSG 968

Search completed: July 3, 2003, 18:23:14
Job time : 13.3157 secs

```

FT DISULFID 461 474 BY SIMILARITY.
FT DISULFID 480 495 BY SIMILARITY.
FT DISULFID 491 504 BY SIMILARITY.
FT DISULFID 506 519 BY SIMILARITY.
FT DISULFID 525 537 BY SIMILARITY.
FT DISULFID 530 546 BY SIMILARITY.
FT DISULFID 551 565 BY SIMILARITY.
FT VARSPLIC 564 712 BY SIMILARITY.
FT ITLKVSAAPHNADPEVNEELQKTTIVGAPNIPAIRAFNL
FT LQGEKRSAYVTLRDSLDGPOTVKQLDLRLSKSKKKNPT
FT YAAALIVVAAKRHNTVHPPLMKIR -> QIADYSCKTV
FT CSTEDTECLGNHTREVLQYFRAVPSIKTIISPIEVSRTVH
FT MGPEVSVDNLDYVGQRHFRIOERNIGIIVOLKPEISGPTV
FT ETIKVNIHTKSRGTGYILAFNEALIEISVSKYFP (IN
FT ISOFORM D)
SQ SEQUENCE 712 AA; 77009 MM; 52CEB8CF8BF296BC5 CRC64;
Query Match 23.7%; Score 599.5; DB 1; Length 712;
Best Local Similarity 27.8%; Pred. No. 2.2e-38;
Matches 155; Conservative 47; Mismatches 153; Indels 203; Gaps 19;
QY 15 LCLPSPG-MQAQCTNFDLDRSGQCLDIDECRIPEACRGMCMVONGYILIPRTN 73
DB 167 LCHDRGGEKEVCCSCRSGLDAPDGMACVDIDECATLMDCLSEOSQCLVTPGSEFKCI- 222
QY 74 PVYRGPSNPSTYSGPYPAAPPLASPNYPTISRPLICRGYQMD-ESNOCVDVDECA 132
DB 223 -----RTLSGGIYAMDESTERCRVDEN 247
QY 133 TDSHCNPOTICINTEGGYTC-----SCYDGYWLLSGQCLDIDE 171
DB 248 LGSHDGCPLYOCRCNTQGSYRCDAKCGDGLONPMTGCTSTICPNGYYPKMGCMNDIDE 307
QY 172 CRGY-----CQO---LCA-----N 183
DB 308 CVYGHNCGAGEECVNTPGSFRCQKGNLCAHGYEVNGATGCEDEVNCOQGVGSMECTIN 367
QY 184 VPGSYSCNPGFTLNE----- 200
DB 368 LPTGYKKCKGPGYEFNDAKKRCEDVDECIRFAGHVCDSABECINTISFECKCKRGFOLA 427
QY 201 -DGRSCDVNECATE-NECVOTCVNTYGSFTRCRDPGYELEBDVGHCSMDSCSF- -S 254
DB 428 SDGRRCDEVNCTTGIAACEQKVINPESYOCICDRGALPBDGKCEBIDECISYIWAGSG 487
QY 255 EPLCOHCVCVNPGRYFPCSPRGYILLDNRSQDINEEHNRHTNONTLOOTCNLGGPRC 314
DB 488 NDLQMGCCINKGSYLQCPRGYKIQPGRTCVDVDECA-MGECAGSDKVCVNTLGSFRC 546
QY 315 IDPIRCEEPYLRISDNR-----CMCPAENPCRCRQDPFTILYRDMDVYSGRSP- 362
DB 547 -HSIDCTYNIHDLNKNRKNRPSACGLPBE--CSNVPLFLTYQFISL-ARAVPISS 600
QY 363 -----ADI--FQMATTTPGAYYIFQIKSGNEGREFFYMRGTGP-ISA 402
DB 601 HRPATITLFKVSAPNADDEVNFELOKTTIVGAPNIPAIRAN-----FLQKGEKRSNA 655
QY 403 TLVATRIKGPREFQDL 420
DB 656 VVTLRLSDSPQVTKQL 673

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fibroblast, and Platelet;
RX MEDLINE=90275601; PubMed=2350783;
RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;
RT "TGF-beta 1 binding protein: a component of the large latent complex
RL of TGF-beta 1 with multiple repeat sequences."
CC Cell 61:1051-1061(1990).
CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE
CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-
CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a
CC long form (AC 014766); are produced by alternative splicing.
CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M34057; AAA61160.1; -.
DR PIR: A35626; A35626.
DR HSP: P00750; TPG.
DR GlycoSiteDB: P22064; -.
DR GeneW: HGNC:6714; LTBPL.
DR MIM: 150390; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR InterPro: IPR002212; Fibrin-assoc.
DR Pfam: PF00008; EGF; 15.
DR Pfam: PF00683; TB; 4.
DR SMART: SM00179; EGF_CA; 13.
DR SMART: SM00001; EGF_Like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; 13.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 11.
DR PROSITE: PS01187; EGF_CA; 15.
KW Growth factor binding; Repeat; EGF-like domain; Hydroxylation; signal;
KW Glycoprotein; Alternative splicing.
FT SIGNAL 20
FT CHAIN 1 1394
FT DOMAIN 300 340
FT REPEAT 348 412
FT DOMAIN 546 587
FT DOMAIN 588 629
FT DOMAIN 630 670
FT DOMAIN 671 710
FT DOMAIN 711 751
FT DOMAIN 752 792
FT DOMAIN 793 833
FT DOMAIN 834 874
FT DOMAIN 875 916
FT DOMAIN 917 958
FT DOMAIN 959 1001
FT REPEAT 1017 1084
FT REPEAT 1097 1139
FT DOMAIN 1190 1262
FT DOMAIN 1140 1180
FT DOMAIN 1294 1334

```

```

RESULT 15
ID LTB_HUMAN STANDARD; PRT; 1394 AA.
AC P22064;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Latent transforming growth factor beta binding protein 15 precursor
DE (transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-
DE 1).
GN LTBPL.

```


[illegible]

Query Match	Best Local Similarity	34.9%	Pred. No. 2.7e-46;	Matches 162;	Conservative 67;	Mismatches 156;	Indels 79;	Gaps 22;
QY	16	CLPSPGNAOACTGNEFDLDRSGGCLIDECRTPEACRGDMCVNONGYLCIPRTNPV	75					
Db	290	CRP-----KLCCKSGFIQD-ALGNCIDIDNEICISAPCPVGGTCINTGSGSYNC-----	336					
QY	76	YRGYSNPSYSPRPSGPPYAPARPLASAPYPTISRLICRFQGM--DESNOCYDVDECATD	134					
Db	337	-----QKNVPR-----CGRGHNLNEBGRVCYDVDECAPR	365					
QY	135	SHOCNPQICINTEGGYCTSCTDGYWL--LRGOCILIDEC-RY--GYCOOLCANYPGSYS	189					
Db	366	AEPGCKGKHGHCNLSRGSFRGCEKAGRYFPGISMTQVINDCQRPERLRLGKHCKENPGRSH	425					
QY	190	CTCNPGFTLNEDGRSCODVNECANEPCVQTCVNTYGSFICRCRDPGELEF--DGVNCSDM	248					
Db	426	CSCSAGFRLSYDGRSCBEDVNEC--LNSPSCQSECCANANYGSGSYOCRCRGOYLSVDVGTCEDI	484					
QY	249	DECSF--SEFLCOHCVNPGTGYTSCSC--PGVILLDDNRSCODINECEARNHTKLOOTC	305					
Db	485	DECALPTEGHCITCRCLINPISFOCSPSSGTRLPARNGRNODDIECVTGHNCINSETC	544					
QY	306	YNLOGGFCIDPINCSEPYLRISDN-----RCM--CPAENPGC-RDOPFTLLYBDM	353					
Db	545	FNIGGSFRCLLS--FCEPEYRRSADTFQREKTDYVACISCRPNDBACVADPNVHTSHYI	603					
QY	354	DVVGGRSV--PADIFOMATP-KRPG--AYIIFOLKSNGEGREFM---RQTRPSAILLV	405					
Db	604	SLPFFREFTRREELIFLRAVPLRYPANADILFTDETNLNDSDFIIRRYEDGATGVYVR	663					
QY	406	MTPRIKGPREIQLDEM-IYVNTYINFGSSVYIRLIYVSQYF	448					
Db	664	QVRIVGEGFYAVLKLKEMTYLVGGVYSHR--NVVNVNHTVSEYWF	705					

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH
 CC BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.
 CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC EMBL: X82494; CA57876.1; -
 CC HSSP: P00736; IAPQ.
 CC Genew: HGNC:3601; FBLN2.
 CC MIM: 135821; -
 CC InterPro: IPR000020; Anaphylatoxin.
 CC InterPro: IPR000152; Asx_hydroxyl.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001881; EGF_Ca.
 CC Pfam: PF00008; EGF_7.
 CC Pfam: PF01821; ANATO; 2.
 CC SMART: SM00104; ANATO; 2.
 CC SMART: SM00179; EGF_Ca; 9.
 CC SMART: SM00001; EGF_Like; 1.
 CC PROSITE: PS00010; ASX_HYDROXYL; 5.
 CC PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 CC PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 CC PROSITE: PS00022; EGF_1; FALSE_NEG.
 CC PROSITE: PS01186; EGF_2; 5.
 CC PROSITE: PS01187; EGF_Ca; 9.
 CC Signal: Glycoprotein. Extracellular matrix; Plasma; EGF-like domain;
 CC Calcium-binding; Repeat.
 CC KW SIGNAL
 CC FT CHAIN 1 27 POTENTIAL.
 CC FT CHAIN 28 1184 FIBULIN-2.
 CC FT DOMAIN 28 444 N.
 CC FT DOMAIN 177 177 SUBDOMAIN NA (CYS-RICH).
 CC FT DOMAIN 178 444 SUBDOMAIN NB (CYS-FREE).
 CC FT DOMAIN 445 480 ANAPHYLATOXIN-LIKE 1.
 CC FT DOMAIN 488 519 ANAPHYLATOXIN-LIKE 2.
 CC FT DOMAIN 521 553 ANAPHYLATOXIN-LIKE 3.
 CC FT DOMAIN 604 645 EGF-LIKE 1, CALCIUM-BINDING.
 CC FT DOMAIN 679 718 EGF-LIKE 2.
 CC FT DOMAIN 719 763 EGF-LIKE 3, CALCIUM-BINDING.
 CC FT DOMAIN 764 809 EGF-LIKE 4, CALCIUM-BINDING.
 CC FT DOMAIN 810 857 EGF-LIKE 5, CALCIUM-BINDING.
 CC FT DOMAIN 858 900 EGF-LIKE 6, CALCIUM-BINDING.
 CC FT DOMAIN 901 942 EGF-LIKE 7, CALCIUM-BINDING.
 CC FT DOMAIN 943 981 EGF-LIKE 8, CALCIUM-BINDING.
 CC FT DOMAIN 982 1024 EGF-LIKE 9, CALCIUM-BINDING.
 CC FT DOMAIN 1025 1069 EGF-LIKE 10, CALCIUM-BINDING.
 CC FT DOMAIN 1070 1184 DOMAIN III.
 CC FT DISULFID 445 472 BY SIMILARITY.
 CC FT DISULFID 446 479 BY SIMILARITY.
 CC FT DISULFID 459 480 BY SIMILARITY.
 CC FT DISULFID 489 518 BY SIMILARITY.
 CC FT DISULFID 502 519 BY SIMILARITY.
 CC FT DISULFID 521 545 BY SIMILARITY.
 CC FT DISULFID 522 552 BY SIMILARITY.
 CC FT DISULFID 535 553 BY SIMILARITY.
 CC FT DISULFID 608 620 BY SIMILARITY.
 CC FT DISULFID 616 629 BY SIMILARITY.
 CC FT DISULFID 629 644 BY SIMILARITY.
 CC FT DISULFID 631 644 BY SIMILARITY.
 CC FT DISULFID 683 693 BY SIMILARITY.
 CC FT DISULFID 689 702 BY SIMILARITY.
 CC FT DISULFID 704 717 BY SIMILARITY.
 CC FT DISULFID 723 736 BY SIMILARITY.
 CC FT DISULFID 730 745 BY SIMILARITY.
 CC FT DISULFID 751 762 BY SIMILARITY.
 CC FT DISULFID 768 781 BY SIMILARITY.
 CC FT DISULFID 775 790 BY SIMILARITY.

FT DISULFID 796 808 BY SIMILARITY.
 FT DISULFID 814 827 BY SIMILARITY.
 FT DISULFID 821 836 BY SIMILARITY.
 FT DISULFID 843 856 BY SIMILARITY.
 FT DISULFID 862 875 BY SIMILARITY.
 FT DISULFID 869 884 BY SIMILARITY.
 FT DISULFID 886 899 BY SIMILARITY.
 FT DISULFID 905 917 BY SIMILARITY.
 FT DISULFID 913 926 BY SIMILARITY.
 FT DISULFID 928 941 BY SIMILARITY.
 FT DISULFID 947 956 BY SIMILARITY.
 FT DISULFID 952 965 BY SIMILARITY.
 FT DISULFID 967 980 BY SIMILARITY.
 FT DISULFID 986 998 BY SIMILARITY.
 FT DISULFID 994 1007 BY SIMILARITY.
 FT DISULFID 1009 1023 BY SIMILARITY.
 FT DISULFID 1029 1042 BY SIMILARITY.
 FT DISULFID 1036 1051 BY SIMILARITY.
 FT DISULFID 1056 1068 BY SIMILARITY.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC...)
 FT CARBOHYD 507 507 N-LINKED (GLCNAC...)
 FT CARBOHYD 1035 1035 N-LINKED (GLCNAC...)
 SO SEQUENCE 1184 AA; 126543 MW; CA48490A5F9ECSD CRC64;
 Query Match 28.0%; Score 709.5; DB 1; Length 1184;
 Best Local Similarity 34.3%; Pred. No. 1,6e-46;
 Matches 146; Conservative 61; Mismatches 156; Indels 63; Gaps 12;
 QY 15 LCLSPRGN---AQAQCTNGFOLDNRSGCGLDDEKRTPEACRGDMCYNGYICIP 70
 DB LCQNTKSGSCARORCMDF-LQDEGNCVINDICTSLSEPCRPSPFCINTGSGYTC-- 836
 QY 71 RTNPKRGPSNPNSTPSGPPAAPPAPNPPTISPLICRFYQ-MDESGCVDV 129
 DB 837 -----QRNPICARGYHASDGAACVDV 860
 QY 130 ECATSDHOCNPQTICINTEGCTSCDTGYW--LEGGCLDDECRG--YCQOLCANV 184
 DB 861 ECETGVHNCGEQYOVHNPGLSGYRCDCAKGFORDAERGCIDVNECAWSPGRLCQHCEMT 920
 QY 185 PGYSCTGNPGFTLNEDRSCDVNECANENPCYQCVNTYSFICRCPGYELEDDYH 244
 DB 921 LGSYRCASASGFLIADGRKCEDVNECEAR--CSGCANVYSYOCYCRQYOLADGHT 979
 QY 245 CSDMECS-FSEFLQHCVCVNPQGYFPCSP-PGYILLDNNSCODINECHRNHTCNQ 302
 DB 980 CTIDDECAOGAGIILCTFRCLNVPGSYOCAPQGYTMANGSCADVDCALGTHNCSEA 1039
 QY 303 QTCYNLOGGFKCIDPIRCEEPYLRISDRCKMPAENPGCRD-----OPTLLYRDMDVY 356
 DB 1040 ETCNHIQSEFKCL-RFECPPNIVQVSKTKG---ERTTCHDPLECONSPPARTTHQLNFO 1094
 QY 357 SGRSVPADIFQMOATTRYGAYVYIFQIKSGNEGRFYMROTPTISATLYMTPIRGPRT 416
 DB 1095 TGLVPAHIFRIGAPAPATGPTIALNITKGNDEGIFGRRLNATYGVYILQAVLEPRDF 1154
 QY 417 QLDDEM 422
 DB 1155 ALDDEM 1160
 RESULT 12
 FBL_MOUSE STANDARD; PRT; 705 AA.
 AC Q08879; Q08879; (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrin-1 precursor (Basement-membrane protein 90) (BM-90).
 GN FBLN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```

FT DOMAIN 400 441 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 442 481 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 482 525 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 526 570 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 33 59 BY SIMILARITY.
FT DISULFID 34 66 BY SIMILARITY.
FT DISULFID 47 67 BY SIMILARITY.
FT DISULFID 76 107 BY SIMILARITY.
FT DISULFID 89 108 BY SIMILARITY.
FT DISULFID 110 134 BY SIMILARITY.
FT DISULFID 111 141 BY SIMILARITY.
FT DISULFID 124 142 BY SIMILARITY.
FT DISULFID 181 191 BY SIMILARITY.
FT DISULFID 187 200 BY SIMILARITY.
FT DISULFID 202 215 BY SIMILARITY.
FT DISULFID 221 234 BY SIMILARITY.
FT DISULFID 228 243 BY SIMILARITY.
FT DISULFID 249 261 BY SIMILARITY.
FT DISULFID 267 280 BY SIMILARITY.
FT DISULFID 274 289 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 313 326 BY SIMILARITY.
FT DISULFID 320 335 BY SIMILARITY.
FT DISULFID 342 355 BY SIMILARITY.
FT DISULFID 361 374 BY SIMILARITY.
FT DISULFID 368 383 BY SIMILARITY.
FT DISULFID 385 398 BY SIMILARITY.
FT DISULFID 404 416 BY SIMILARITY.
FT DISULFID 412 425 BY SIMILARITY.
FT DISULFID 427 440 BY SIMILARITY.
FT DISULFID 446 455 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 466 480 BY SIMILARITY.
FT DISULFID 486 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 530 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 557 569 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 684 AA: 75623 MW: 1638D7A094739199 CRC64;

Query Match 30.7%; Score 778.5; DB 1; Length 684;
Best Local Similarity 35.9%; Pred. No. 5,3e-52;
Matches 162; Conservative 66; Mismatches 154; Indels 69; Gaps 16;

```

```

OY 359 RSPVPAIFOMQATTRIPGAVYIFQIKSGNEGRFYMRQGPISATLVWTRPKPREIOL 418
DB 598 IQVPTDIFRMPGPNAPVGDGKILLISISGNEGFTTKVNNHSGIVWQRIPTPRULL 657
OY 419 DLEMI-----TWTYVINFRRSSVIRLRIYYS 444
DB 658 TIOMQLTRHGVNTFTI-----AKLFEVVS 681

RESULT 10
FBL2_MOUSE
ID FBL2_MOUSE STANDARD; PRT; 1221 AA.
AC P37889; Q9WU12.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fbln2,
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast.
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding."
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99337686; PubMed=10406956;
RA Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization."
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC CALCULON DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms: 1 (shown here) and
CC 2/EGF-less; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X75285; CAA53040.1; -
DR EMBL; AF135253; AAD34456.1; -
DR EMBL; AF135253; AAD34456.1; JOINED.
DR EMBL; AF135240; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135242; AAD34456.1; JOINED.
DR EMBL; AF135243; AAD34456.1; JOINED.
DR EMBL; AF135244; AAD34456.1; JOINED.
DR EMBL; AF135245; AAD34456.1; JOINED.
DR EMBL; AF135246; AAD34456.1; JOINED.
DR EMBL; AF135247; AAD34456.1; JOINED.
DR EMBL; AF135248; AAD34456.1; JOINED.
DR EMBL; AF135249; AAD34456.1; JOINED.
DR EMBL; AF135250; AAD34456.1; JOINED.
DR EMBL; AF135251; AAD34456.1; JOINED.
DR EMBL; AF135252; AAD34456.1; JOINED.

```

DR Pfam: PF00008; EGF_3.
 DR SMART: SM00179; EGF_CA: 4.
 DR SMART: SM00001; EGF_Like: 2.
 DR PROSITE: PS00010; ASX_HYDROXYL: 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS01187; EGF_CA: 6.
 DR Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal;
 KW Disease mutation; Polymorphism; Alternative splicing.
 FT SIGNAL 1 17
 FT CHAIN 18 493
 FT
 FT DOMAIN 26 71
 FT 173 213
 FT DOMAIN 214 253
 FT 254 293
 FT DOMAIN 294 333
 FT 334 378
 FT DOMAIN 378 493
 FT 177 199
 FT DISULFID 184 199
 FT 201 212
 FT DISULFID 218 228
 FT 224 237
 FT DISULFID 239 252
 FT 258 268
 FT DISULFID 264 277
 FT 279 292
 FT DISULFID 298 309
 FT 318 332
 FT DISULFID 330 332
 FT 338 350
 FT DISULFID 344 359
 FT 365 377
 FT CARBOHYD 249 249
 FT 249 249
 FT VARPPLIC 1 8
 FT 58 58
 FT VARPPLIC 106 106
 FT 220 220
 FT VARIANT 345 345
 FT 345 345
 FT
 FT SEQUENCE 493 AA: 54640 MW: 128CA5ED140DF414 CRC64;
 SO
 Query Match 43.5%; Score 1101; DB 1; Length 493;
 Best Local Similarity 42.0%; Pred. No. 1,2e-76;
 Matches 208; Conservative 69; Mismatches 162; Indels 56; Gaps 8;
 6 RLVYITIALCLPSPENAO-----AQTNGFDLDRSGGCLDDECTIPRACRGDMKV 60
 3 KALFTMLTLALVKSQDTEETIYTOCTDQEMDPVROCKDIDCDIVPCKGKMKCV 62
 61 NONGYLCIPRNPV---RGPSNPY-----STPSGPRPA----- 95
 63 NHYGILCLPKTAQIIIVNNEQPOEQTQPAEGTSGATTGVAASMATSGVLPGGGFVASA 122
 96 -----APPLSAPNPITISRPICRFGYOMDSNOCVDYDECATD 134
 123 AAVAPREMGQTGNFPIRRNPADPRLPSNP--SHRQCAAGYQSHHNQODIDECAG 180
 135 SHQCPPTQICITEGGYTCSDTDGYMLLESGCLDIDECRY-GYCOQLCANVPSGYSCTCN 193
 181 THNCRADYVCIMLRGSAFCQCPGQYKRGECVADIDECCTIPYCHORCVMPGSGFYCCS 240
 194 PGFTLNEDRGSCADVNECATENPCVOTCVATVSGFSCRCRCPGVELERDGYHSCMDSCF 253
 241 PGFOLANNNTYCVDINECDASNOCACQCVNLLSFCQCKGVELSSDRLNCEDIDECRT 300
 254 SEFLCOHCEVNOPGYFYSCSPRYILLDNRSQCODINECHRNHCTNMLQOTCYMLGGFK 313
 301 SSYLQYQCVNRPGRKSCMCPOGYVY-NSRTQDIDNECTIN-CREDEMCNMYHGGFR 358
 314 CIDPRCEPRYLRISDNRCMCRAENPCRDQPTILYRDMVVSGRSVPADIFOMQATTR 373

DB 359 CYPNRCQDPYILTPENRCVCPVSNAMCRELPQSIYKYNISRSRVSPSDIFOIQAATTI 418
 QY 374 YGAYYIFQIKSGNENGEFMRGTGPRISATLVMTNRPJKPREIOLDLMTLVNVIYNRG 433
 DB 419 YANTINTFRKSGNENGEFYLRTQSPVSNALVLKVSIGPREHLYDLEMLTVSSIGTERT 478
 QY 434 SSYVRLRIYVSQYPF 448
 DB 479 SSYVRLRIYVGPSPF 493
 RESULT 9
 FBLLCHICK STANDARD; PRT: 684 AA.
 ID FBLLCHICK 073775;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibulin-1 precursor.
 GN FBLLN1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120531; PubMed=9923656.
 RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
 RT "Identification of chicken and C. elegans fibulin-1 homologs and
 RT characterization of the C. elegans fibulin-1 gene."
 RL Matrix Biol. 17:635-646(1998).
 CC -1 SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -1 SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1 SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF051400; AAC05388.1; .
 CC HSP: P00742; IHCG.
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; Asx_hydroxyl.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_6.
 DR Pfam: PRO1821; ANATO_2.
 DR SMART: SM00104; ANATO_3.
 DR SMART: SM00179; EGF_CA: 8.
 DR SMART: SM00001; EGF_Like: 1.
 DR PROSITE: PS00010; ASX_HYDROXYL: 5.
 DR PROSITE: PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE: PS01178; ANAPHYLATOXIN_2; 2.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA: 8.
 KW Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;
 KW Calcium-binding.
 FT SIGNAL 1 25
 FT CHAIN 26 684
 FT DOMAIN 33 74
 FT 75 109
 FT DOMAIN 110 142
 FT 143 177
 FT DOMAIN 177 216
 FT 217 262
 FT DOMAIN 263 308
 FT 309 356
 FT DOMAIN 357 399
 FT
 FT POTENTIAL.
 FT FIBULIN-1.
 FT ANAPHYLATOXIN-Like 1.
 FT ANAPHYLATOXIN-Like 2.
 FT ANAPHYLATOXIN-Like 3.
 FT EGF-Like 1.
 FT EGF-Like 2.
 FT EGF-Like 3.
 FT EGF-Like 4.
 FT EGF-Like 5.
 FT EGF-Like 6.
 FT EGF-Like 7.
 FT EGF-Like 8.
 FT EGF-Like 9.
 FT EGF-Like 10.
 FT EGF-Like 11.
 FT EGF-Like 12.
 FT EGF-Like 13.
 FT EGF-Like 14.
 FT EGF-Like 15.
 FT EGF-Like 16.
 FT EGF-Like 17.
 FT EGF-Like 18.
 FT EGF-Like 19.
 FT EGF-Like 20.
 FT EGF-Like 21.
 FT EGF-Like 22.
 FT EGF-Like 23.
 FT EGF-Like 24.
 FT EGF-Like 25.
 FT EGF-Like 26.
 FT EGF-Like 27.
 FT EGF-Like 28.
 FT EGF-Like 29.
 FT EGF-Like 30.
 FT EGF-Like 31.
 FT EGF-Like 32.
 FT EGF-Like 33.
 FT EGF-Like 34.
 FT EGF-Like 35.
 FT EGF-Like 36.
 FT EGF-Like 37.
 FT EGF-Like 38.
 FT EGF-Like 39.
 FT EGF-Like 40.
 FT EGF-Like 41.
 FT EGF-Like 42.
 FT EGF-Like 43.
 FT EGF-Like 44.
 FT EGF-Like 45.
 FT EGF-Like 46.
 FT EGF-Like 47.
 FT EGF-Like 48.
 FT EGF-Like 49.
 FT EGF-Like 50.
 FT EGF-Like 51.
 FT EGF-Like 52.
 FT EGF-Like 53.
 FT EGF-Like 54.
 FT EGF-Like 55.
 FT EGF-Like 56.
 FT EGF-Like 57.
 FT EGF-Like 58.
 FT EGF-Like 59.
 FT EGF-Like 60.
 FT EGF-Like 61.
 FT EGF-Like 62.
 FT EGF-Like 63.
 FT EGF-Like 64.
 FT EGF-Like 65.
 FT EGF-Like 66.
 FT EGF-Like 67.
 FT EGF-Like 68.
 FT EGF-Like 69.
 FT EGF-Like 70.
 FT EGF-Like 71.
 FT EGF-Like 72.
 FT EGF-Like 73.
 FT EGF-Like 74.
 FT EGF-Like 75.
 FT EGF-Like 76.
 FT EGF-Like 77.
 FT EGF-Like 78.
 FT EGF-Like 79.
 FT EGF-Like 80.
 FT EGF-Like 81.
 FT EGF-Like 82.
 FT EGF-Like 83.
 FT EGF-Like 84.
 FT EGF-Like 85.
 FT EGF-Like 86.
 FT EGF-Like 87.
 FT EGF-Like 88.
 FT EGF-Like 89.
 FT EGF-Like 90.
 FT EGF-Like 91.
 FT EGF-Like 92.
 FT EGF-Like 93.
 FT EGF-Like 94.
 FT EGF-Like 95.
 FT EGF-Like 96.
 FT EGF-Like 97.
 FT EGF-Like 98.
 FT EGF-Like 99.
 FT EGF-Like 100.
 FT EGF-Like 101.
 FT EGF-Like 102.
 FT EGF-Like 103.
 FT EGF-Like 104.
 FT EGF-Like 105.
 FT EGF-Like 106.
 FT EGF-Like 107.
 FT EGF-Like 108.
 FT EGF-Like 109.
 FT EGF-Like 110.
 FT EGF-Like 111.
 FT EGF-Like 112.
 FT EGF-Like 113.
 FT EGF-Like 114.
 FT EGF-Like 115.
 FT EGF-Like 116.
 FT EGF-Like 117.
 FT EGF-Like 118.
 FT EGF-Like 119.
 FT EGF-Like 120.
 FT EGF-Like 121.
 FT EGF-Like 122.
 FT EGF-Like 123.
 FT EGF-Like 124.
 FT EGF-Like 125.
 FT EGF-Like 126.
 FT EGF-Like 127.
 FT EGF-Like 128.
 FT EGF-Like 129.
 FT EGF-Like 130.
 FT EGF-Like 131.
 FT EGF-Like 132.
 FT EGF-Like 133.
 FT EGF-Like 134.
 FT EGF-Like 135.
 FT EGF-Like 136.
 FT EGF-Like 137.
 FT EGF-Like 138.
 FT EGF-Like 139.
 FT EGF-Like 140.
 FT EGF-Like 141.
 FT EGF-Like 142.
 FT EGF-Like 143.
 FT EGF-Like 144.
 FT EGF-Like 145.
 FT EGF-Like 146.
 FT EGF-Like 147.
 FT EGF-Like 148.
 FT EGF-Like 149.
 FT EGF-Like 150.
 FT EGF-Like 151.
 FT EGF-Like 152.
 FT EGF-Like 153.
 FT EGF-Like 154.
 FT EGF-Like 155.
 FT EGF-Like 156.
 FT EGF-Like 157.
 FT EGF-Like 158.
 FT EGF-Like 159.
 FT EGF-Like 160.
 FT EGF-Like 161.
 FT EGF-Like 162.
 FT EGF-Like 163.
 FT EGF-Like 164.
 FT EGF-Like 165.
 FT EGF-Like 166.
 FT EGF-Like 167.
 FT EGF-Like 168.
 FT EGF-Like 169.
 FT EGF-Like 170.
 FT EGF-Like 171.
 FT EGF-Like 172.
 FT EGF-Like 173.
 FT EGF-Like 174.
 FT EGF-Like 175.
 FT EGF-Like 176.
 FT EGF-Like 177.
 FT EGF-Like 178.
 FT EGF-Like 179.
 FT EGF-Like 180.
 FT EGF-Like 181.
 FT EGF-Like 182.
 FT EGF-Like 183.
 FT EGF-Like 184.
 FT EGF-Like 185.
 FT EGF-Like 186.
 FT EGF-Like 187.
 FT EGF-Like 188.
 FT EGF-Like 189.
 FT EGF-Like 190.
 FT EGF-Like 191.
 FT EGF-Like 192.
 FT EGF-Like 193.
 FT EGF-Like 194.
 FT EGF-Like 195.
 FT EGF-Like 196.
 FT EGF-Like 197.
 FT EGF-Like 198.
 FT EGF-Like 199.
 FT EGF-Like 200.
 FT EGF-Like 201.
 FT EGF-Like 202.
 FT EGF-Like 203.
 FT EGF-Like 204.
 FT EGF-Like 205.
 FT EGF-Like 206.
 FT EGF-Like 207.
 FT EGF-Like 208.
 FT EGF-Like 209.
 FT EGF-Like 210.
 FT EGF-Like 211.
 FT EGF-Like 212.
 FT EGF-Like 213.
 FT EGF-Like 214.
 FT EGF-Like 215.
 FT EGF-Like 216.
 FT EGF-Like 217.
 FT EGF-Like 218.
 FT EGF-Like 219.
 FT EGF-Like 220.
 FT EGF-Like 221.
 FT EGF-Like 222.
 FT EGF-Like 223.
 FT EGF-Like 224.
 FT EGF-Like 225.
 FT EGF-Like 226.
 FT EGF-Like 227.
 FT EGF-Like 228.
 FT EGF-Like 229.
 FT EGF-Like 230.
 FT EGF-Like 231.
 FT EGF-Like 232.
 FT EGF-Like 233.
 FT EGF-Like 234.
 FT EGF-Like 235.
 FT EGF-Like 236.
 FT EGF-Like 237.
 FT EGF-Like 238.
 FT EGF-Like 239.
 FT EGF-Like 240.
 FT EGF-Like 241.
 FT EGF-Like 242.
 FT EGF-Like 243.
 FT EGF-Like 244.
 FT EGF-Like 245.
 FT EGF-Like 246.
 FT EGF-Like 247.
 FT EGF-Like 248.
 FT EGF-Like 249.
 FT EGF-Like 250.
 FT EGF-Like 251.
 FT EGF-Like 252.
 FT EGF-Like 253.
 FT EGF-Like 254.
 FT EGF-Like 255.
 FT EGF-Like 256.
 FT EGF-Like 257.
 FT EGF-Like 258.
 FT EGF-Like 259.
 FT EGF-Like 260.
 FT EGF-Like 261.
 FT EGF-Like 262.
 FT EGF-Like 263.
 FT EGF-Like 264.
 FT EGF-Like 265.
 FT EGF-Like 266.
 FT EGF-Like 267.
 FT EGF-Like 268.
 FT EGF-Like 269.
 FT EGF-Like 270.
 FT EGF-Like 271.
 FT EGF-Like 272.
 FT EGF-Like 273.
 FT EGF-Like 274.
 FT EGF-Like 275.
 FT EGF-Like 276.
 FT EGF-Like 277.
 FT EGF-Like 278.
 FT EGF-Like 279.
 FT EGF-Like 280.
 FT EGF-Like 281.
 FT EGF-Like 282.
 FT EGF-Like 283.
 FT EGF-Like 284.
 FT EGF-Like 285.
 FT EGF-Like 286.
 FT EGF-Like 287.
 FT EGF-Like 288.
 FT EGF-Like 289.
 FT EGF-Like 290.
 FT EGF-Like 291.
 FT EGF-Like 292.
 FT EGF-Like 293.
 FT EGF-Like 294.
 FT EGF-Like 295.
 FT EGF-Like 296.
 FT EGF-Like 297.
 FT EGF-Like 298.
 FT EGF-Like 299.
 FT EGF-Like 300.
 FT EGF-Like 301.
 FT EGF-Like 302.
 FT EGF-Like 303.
 FT EGF-Like 304.
 FT EGF-Like 305.
 FT EGF-Like 306.
 FT EGF-Like 307.
 FT EGF-Like 308.
 FT EGF-Like 309.
 FT EGF-Like 310.
 FT EGF-Like 311.
 FT EGF-Like 312.
 FT EGF-Like 313.
 FT EGF-Like 314.
 FT EGF-Like 315.
 FT EGF-Like 316.
 FT EGF-Like 317.
 FT EGF-Like 318.
 FT EGF-Like 319.
 FT EGF-Like 320.
 FT EGF-Like 321.
 FT EGF-Like 322.
 FT EGF-Like 323.
 FT EGF-Like 324.
 FT EGF-Like 325.
 FT EGF-Like 326.
 FT EGF-Like 327.
 FT EGF-Like 328.
 FT EGF-Like 329.
 FT EGF-Like 330.
 FT EGF-Like 331.
 FT EGF-Like 332.
 FT EGF-Like 333.
 FT EGF-Like 334.
 FT EGF-Like 335.
 FT EGF-Like 336.
 FT EGF-Like 337.
 FT EGF-Like 338.
 FT EGF-Like 339.
 FT EGF-Like 340.
 FT EGF-Like 341.
 FT EGF-Like 342.
 FT EGF-Like 343.
 FT EGF-Like 344.
 FT EGF-Like 345.
 FT EGF-Like 346.
 FT EGF-Like 347.
 FT EGF-Like 348.
 FT EGF-Like 349.
 FT EGF-Like 350.
 FT EGF-Like 351.
 FT EGF-Like 352.
 FT EGF-Like 353.
 FT EGF-Like 354.
 FT EGF-Like 355.
 FT EGF-Like 356.
 FT EGF-Like 357.
 FT EGF-Like 358.
 FT EGF-Like 359.
 FT EGF-Like 360.
 FT EGF-Like 361.
 FT EGF-Like 362.
 FT EGF-Like 363.
 FT EGF-Like 364.
 FT EGF-Like 365.
 FT EGF-Like 366.
 FT EGF-Like 367.
 FT EGF-Like 368.
 FT EGF-Like 369.
 FT EGF-Like 370.
 FT EGF-Like 371.
 FT EGF-Like 372.
 FT EGF-Like 373.
 FT EGF-Like 374.
 FT EGF-Like 375.
 FT EGF-Like 376.
 FT EGF-Like 377.
 FT EGF-Like 378.
 FT EGF-Like 379.
 FT EGF-Like 380.
 FT EGF-Like 381.
 FT EGF-Like 382.
 FT EGF-Like 383.
 FT EGF-Like 384.
 FT EGF-Like 385.
 FT EGF-Like 386.
 FT EGF-Like 387.
 FT EGF-Like 388.
 FT EGF-Like 389.
 FT EGF-Like 390.
 FT EGF-Like 391.
 FT EGF-Like 392.
 FT EGF-Like 393.
 FT EGF-Like 394.
 FT EGF-Like 395.
 FT EGF-Like 396.
 FT EGF-Like 397.
 FT EGF-Like 398.
 FT EGF-Like 399.
 FT EGF-Like 400.
 FT EGF-Like 401.
 FT EGF-Like 402.
 FT EGF-Like 403.
 FT EGF-Like 404.
 FT EGF-Like 405.
 FT EGF-Like 406.
 FT EGF-Like 407.
 FT EGF-Like 408.
 FT EGF-Like 409.
 FT EGF-Like 410.
 FT EGF-Like 411.
 FT EGF-Like 412.
 FT EGF-Like 413.
 FT EGF-Like 414.
 FT EGF-Like 415.
 FT EGF-Like 416.
 FT EGF-Like 417.
 FT EGF-Like 418.
 FT EGF-Like 419.
 FT EGF-Like 420.
 FT EGF-Like 421.
 FT EGF-Like 422.
 FT EGF-Like 423.
 FT EGF-Like 424.
 FT EGF-Like 425.
 FT EGF-Like 426.
 FT EGF-Like 427.
 FT EGF-Like 428.
 FT EGF-Like 429.
 FT EGF-Like 430.
 FT EGF-Like 431.
 FT EGF-Like 432.
 FT EGF-Like 433.
 FT EGF-Like 434.
 FT EGF-Like 435.
 FT EGF-Like 436.
 FT EGF-Like 437.
 FT EGF-Like 438.
 FT EGF-Like 439.
 FT EGF-Like 440.
 FT EGF-Like 441.
 FT EGF-Like 442.
 FT EGF-Like 443.
 FT EGF-Like 444.
 FT EGF-Like 445.
 FT EGF-Like 446.
 FT EGF-Like 447.
 FT EGF-Like 448.
 FT EGF-Like 449.
 FT EGF-Like 450.
 FT EGF-Like 451.
 FT EGF-Like 452.
 FT EGF-Like 453.
 FT EGF-Like 454.
 FT EGF-Like 455.
 FT EGF-Like 456.
 FT EGF-Like 457.
 FT EGF-Like 458.
 FT EGF-Like 459.
 FT EGF-Like 460.
 FT EGF-Like 461.
 FT EGF-Like 462.
 FT EGF-Like 463.
 FT EGF-Like 464.
 FT EGF-Like 465.
 FT EGF-Like 466.
 FT EGF-Like 467.
 FT EGF-Like 468.
 FT EGF-Like 469.
 FT EGF-Like 470.
 FT EGF-Like 471.
 FT EGF-Like 472.
 FT EGF-Like 473.
 FT EGF-Like 474.
 FT EGF-Like 475.
 FT EGF-Like 476.
 FT EGF-Like 477.
 FT EGF-Like 478.
 FT EGF-Like 479.
 FT EGF-Like 480.
 FT EGF-Like 481.
 FT EGF-Like 482.
 FT EGF-Like 483.
 FT EGF-Like 484.
 FT EGF-Like 485.
 FT EGF-Like 486.
 FT EGF-Like 487.
 FT EGF-Like 488.
 FT EGF-Like 489.
 FT EGF-Like 490.
 FT EGF-Like 491.
 FT EGF-Like 492.
 FT EGF-Like 493.
 FT EGF-Like 494.
 FT EGF-Like 495.
 FT EGF-Like 496.
 FT EGF-Like 497.
 FT EGF-Like 498.
 FT EGF-Like 499.
 FT EGF-Like 500.
 FT EGF-Like 501.
 FT EGF-Like 502.
 FT EGF-Like 503.
 FT EGF-Like 504.
 FT EGF-Like 505.
 FT EGF-Like 506.
 FT EGF-Like 507.
 FT EGF-Like 508.
 FT EGF-Like 509.
 FT EGF-Like 510.
 FT EGF-Like 511.
 FT EGF-Like 512.
 FT EGF-Like 513.
 FT EGF-Like 514.
 FT EGF-Like 515.
 FT EGF-Like 516.
 FT EGF-Like 517.
 FT EGF-Like 518.
 FT EGF-Like 519.
 FT EGF-Like 520.
 FT EGF-Like 521.
 FT EGF-Like 522.
 FT EGF-Like 523.
 FT EGF-Like 524.
 FT EGF-Like 525.
 FT EGF-Like 526.
 FT EGF-Like 527.
 FT EGF-Like 528.
 FT EGF-Like 529.
 FT EGF-Like 530.
 FT EGF-Like 531.
 FT EGF-Like 532.
 FT EGF-Like 533.
 FT EGF-Like 534.
 FT EGF-Like 535.
 FT EGF-Like 536.
 FT EGF-Like 537.
 FT EGF-Like 538.
 FT EGF-Like 539.
 FT EGF-Like 540.
 FT EGF-Like 541.
 FT EGF-Like 542.
 FT EGF-Like 543.
 FT EGF-Like 544.
 FT EGF-Like 545.
 FT EGF-Like 546.
 FT EGF-Like 547.
 FT EGF-Like 548.
 FT EGF-Like 549.
 FT EGF-Like 550.
 FT EGF-Like 551.
 FT EGF-Like 552.
 FT EGF-Like 553.
 FT EGF-Like 554.
 FT EGF-Like 555.
 FT EGF-Like 556.
 FT EGF-Like 557.
 FT EGF-Like 558.
 FT EGF-Like 559.
 FT EGF-Like 560.
 FT EGF-Like 561.
 FT EGF-Like 562.
 FT EGF-Like 563.
 FT EGF-Like 564.
 FT EGF-Like 565.
 FT EGF-Like 566.
 FT EGF-Like 567.
 FT EGF-Like 568.
 FT EGF-Like 569.
 FT EGF-Like 570.
 FT EGF-Like 571.
 FT EGF-Like 572.
 FT EGF-Like 573.
 FT EGF-Like 574.
 FT EGF-Like 575.
 FT EGF-Like 576.
 FT EGF-Like 577.
 FT EGF-Like 578.
 FT EGF-Like 579.
 FT EGF-Like 580.
 FT EGF-Like 581.
 FT EGF-Like 582.
 FT EGF-Like 583.
 FT EGF-Like 584.
 FT EGF-Like 585.
 FT EGF-Like 586.
 FT EGF-Like 587.
 FT EGF-Like 588.
 FT EGF-Like 589.
 FT EGF-Like 590.
 FT EGF-Like 591.
 FT EGF-Like 592.
 FT EGF-Like 593.
 FT EGF-Like 594.
 FT EGF-Like 595.
 FT EGF-Like 596.
 FT EGF-Like 597.
 FT EGF-Like 598.
 FT EGF-Like 599.
 FT EGF-Like 600.
 FT EGF-Like 601.
 FT EGF-Like 602.
 FT EGF-Like 603.
 FT EGF-Like 604.
 FT EGF-Like 605.
 FT EGF-Like 606.
 FT EGF-Like 607.
 FT EGF-Like 608.
 FT EGF-Like 609.
 FT EGF-Like 610.
 FT EGF-Like 611.
 FT EGF-Like 612.
 FT EGF-Like 613.
 FT EGF-Like 614.
 FT EGF-Like 615.
 FT EGF-Like 616.
 FT EGF-Like 617.
 FT EGF-Like 618.
 FT EGF-Like 619.
 FT EGF-Like 620.
 FT EGF-Like 621.
 FT EGF-Like 622.
 FT EGF-Like 623.
 FT EGF-Like 624.
 FT EGF-Like 625.
 FT EGF-Like 626.
 FT EGF-Like 627.
 FT EGF-Like 628.
 FT EGF-Like 629.
 FT EGF-Like 630.
 FT EGF-Like 631.
 FT EGF-Like 632.
 FT EGF-Like 633.
 FT EGF-Like 634.
 FT EGF-Like 635.
 FT EGF-Like 636.
 FT EGF-Like 637.
 FT EGF-Like 638.
 FT EGF-Like 639.
 FT EGF-Like 640.
 FT EGF-Like 641.
 FT EGF-Like 642.
 FT EGF-Like 643.
 FT EGF-Like 644.
 FT EGF-Like 645.
 FT EGF-Like 646.
 FT EGF-Like 647.
 FT EGF-Like 648.
 FT EGF-Like 649.
 FT EGF-Like 650.
 FT EGF-Like 651.
 FT EGF-Like 652.
 FT EGF-Like 653.
 FT EGF-Like 654.
 FT EGF-Like 655.
 FT EGF-Like 656.
 FT EGF-Like 657.
 FT EGF-Like 658.
 FT EGF-Like 659.
 FT EGF-Like 660.
 FT EGF-Like 661.
 FT EGF-Like 662.
 FT EGF-Like 663.
 FT EGF-Like 664.
 FT EGF-Like 665.
 FT EGF-Like 666.
 FT EGF-Like 667.
 FT EGF-Like 668.
 FT EGF-Like 669.
 FT EGF-Like 670.
 FT EGF-Like 671.
 FT EGF-Like 672.
 FT EGF-Like 673.
 FT EGF-Like 674.
 FT EGF-Like 675.
 FT EGF-Like 676.
 FT EGF-Like 677.
 FT EGF-Like 678.
 FT EGF-Like 679.
 FT EGF-Like 68

```

DR PROSITE: PS00010; ASX-HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA; 6.
FT Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 493
FT DOMAIN 26 71
FT DOMAIN 173 213
FT DOMAIN 214 253
FT DOMAIN 254 293
FT DOMAIN 294 333
FT DOMAIN 334 378
FT DISULFID 177 190
FT DISULFID 184 199
FT DISULFID 201 212
FT DISULFID 218 228
FT DISULFID 224 237
FT DISULFID 239 252
FT DISULFID 258 268
FT DISULFID 264 277
FT DISULFID 278 292
FT DISULFID 305 318
FT DISULFID 320 332
FT DISULFID 338 350
FT DISULFID 344 359
FT DISULFID 365 377
SQ SEQUENCE 493 AA; 54596 MW; 22DBFD70BACF1CA5 CRC64;

Query Match 43.8%; Score 1109.5; DB 1; Length 493;
Best Local Similarity 42.3%; Pred. No. 2,7e-77;
Matches 206; Conservative 73; Mismatches 161; Indels 47; Gaps 5;

7 ILTVLLALCLPSPNAOACTNGEDLDROSCDIDDCRTIPACRGDMACVNOGGY 66
9 MTLTALVSVQVEETITTYTQCTDGYEMDPVROCKDIDCDIDIPDACKGMKCVNHGGY 68
67 LCIRTNVNVGRPYSNPYSTPS-----GPPAAPPASLP 102
69 LCLPRTAQIIYNNRPOQETPRAEASSGATGTIARSMATGVTIPGGGFIASATVAGP 128
103 NYPT-----ISRLPICRGYQMDSESNOCVDVDECATDSHOCNPTQ 142
129 EVQYGRNNFVIRNPADPQRIIPSNDSHRIQCAAGYEGSHVNCQDIDECTSGHNRLQ 188
143 ICINTEGGYTSCDGYWLLBEGCCDIDECRT-GYCOQLCANVPGSYSTCNDPFTLND 201
189 VCIILRGSTFCHCLPGYOKRBOQCDIDECVPPYCHOCQVMTPGSFYCOCPNGFOLAN 248
202 GRSCQDVNECATENPCVOTCVNTYGSFTCRDPGYLEEDGYHCSMDSCSFSEFLCOH 261
249 NYTCVDINECDASNOCAQCYNIIGSFICOCNOGVELSDRINCEDIDECRTSSYLCOQ 308
262 CVNPGTYFFCCCPGYILLDDNRSCQDINECHRNHTCNLQOCTYNLAGFCIDIRCE 321
309 CVNPGTYFFCCCPGYILLDDNRSCQDINECHRNHTCNLQOCTYNLAGFCIDIRCE 366
322 EPLRLIDNRMCAPENRGRDOPFTILYRDMVDYSGRSPADIDOMATTPPGAYTF 381
367 DPVYLTSNRCYCVSNMCMRDVPOSIVYKYNINISDSVSDIOLQATTTIYANTINF 426
382 QISGNGREFYMQGTGISATLVMTRPDKPREIQLDLEMTIVTVINFRGSSVIRLRI 441
427 RISGNGREFYMLKOTSPVSMALVKSLSLTPREHIVLEMLTVSSIGFRRSSVLRLLI 486
442 YVSOYPF 448
487 IVGPFSF 493

```

RESULT 8

```

FBL3_HUMAN
ID FBL3_HUMAN STANDARD; PRT; 493 AA.
AC Q12805;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE EGF-containing fibulin-like extracellular matrix protein 1 precursor
DE (Fibulin-3) (FibL-3) (Fibillin-like protein) (Extracellular protein
DE S1-5).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI-TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND POSSIBLE ALTERNATIVE SPLICING.
RC TISSUE=Skin;
RX MEDLINE=95097983; PubMed=7799918;
RA Lecka-Czerwik B., Lumpkin C.K. Jr., Goldstein S.;
RT "An overexpressed gene transcript in senescent and quiescent human
RT fibroblasts encoding a novel protein in the epidermal growth factor-
RT like repeat family stimulates DNA synthesis."
RL Mol. Cell. Biol. 15:120-128(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001163; PubMed=8812496;
RA Ikegawa S., Toda T., Okui K., Nakamura Y.;
RT "Structure and chromosomal assignment of the human S1-5 gene (FBL3)
RT that is highly homologous to fibrillin."
RL Genomics 35:590-592(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20068041; PubMed=10601734;
RA Giltay R., Timpl R., Kostka G.;
RT "Sequence, recombinant expression and tissue localization of two novel
RT extracellular matrix proteins, fibulin-3 and fibulin-4."
RL Matrix Biol. 18:469-480(1999).
RN [4]
RP VARIANT DHRD/MLVT TRP-345, AND VARIANT PHE-220.
RX MEDLINE=99295941; PubMed=10369267;
RA Stone E.M., Lotery A.J., Munier F.L., Heon E., Pignet B., Guymer R.H.,
RA Vandenberg K., Cousin P., Nishimura D., Swiderski R.E., Silvestri G.,
RA Mackey D.A., Hageman G.S., Bird A.C., Sheffield V.C.,
RA Schorderet D.F.;
RT "A single EFEMP1 mutation associated with both malattia Leventinese
RT and Doyme honeycomb retinal dystrophy."
RL Nat. Genet. 22:199-202(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; may
CC be produced by alternative splicing.
CC -1- DISEASE: DEFECTS IN EFEMP1 ARE A CAUSE OF DOYNE HONEYCOMB RETINAL
CC DYSTROPHY (DHRD) ALSO KNOWN AS MALATTIA LEVENTINESE (MLVT OR ML),
CC AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY YELLOW-WHITE
CC DEPOSITS KNOWN AS DRUSEN THAT ACCUMULATE BENEATH THE RETINAL
CC PIGMENT EPITHELIUM.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03877; AAA65590.1; -.
DR HSSP: P35535; 1EMN.
DR Genew: HGNC:3218; EFEMP1.
DR MIM: 601548; -.
DR MIM: 126600; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.

```


16-OCT-2001 (Rel. 40, last annotation update)
 EGF-containing fibulin-4-like extracellular matrix protein 2 precursor
 (Fibulin-4) (FIBL-4) (UPH1 protein).
 EFEMP2 OR FBLN4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID:9606;
 [1]
 SEQUENCE FROM N.A.
 RC Tissue-Melanoma;
 MEDLINE:20068041; PubMed:10601734;
 RA Gilay R., Timpl R., Koska G.;
 "Sequence, recombinant expression and tissue localization of two novel
 extracellular matrix proteins, fibulin-3 and fibulin-4.";
 Matrix Biol. 18:469-480(1999).
 [2]
 SEQUENCE FROM N.A.
 RP Zemel R., Shaul Y.;
 Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 SEQUENCE FROM N.A.
 MEDLINE:20435063; PubMed:10982184;
 RA Katsanis N., Venable S., Smith J.R., Lupski J.R.;
 "Isolation of a paralog of the Doyme honeycomb retinal dystrophy gene
 from the multiple retinopathy critical region on 11q13.";
 Hum. Genet. 106:66-72(2000).
 [4]
 SEQUENCE FROM N.A.
 RC Tissue-Brain;
 Strausberg R.;
 Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SUBCELLULAR LOCATION: Secreted.
 CC -1 SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AJ132819; CA10791.2; -
 DR EMBL; AF093119; AAC62108.1; -
 DR EMBL; AF109121; AAF65188.1; -
 DR EMBL; BC010456; AAH10456.1; -
 DR HSSP; P35555; IEMN.
 DR Gene; HGNC:3219; EFEMP2.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR001491; Thrombomodulin.
 DR Pfam; PF00008; EGF; 4.
 DR PRINTS; PR00907; THROMBOMODULN.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; signal.
 FT SIGNAL 1 25
 FT CHAIN 26 443
 FT MATRIX PROTEIN 2.
 FT DOMAIN 36 81
 FT DOMAIN 123 163
 FT DOMAIN 164 202
 FT DOMAIN 203 242
 FT DOMAIN 243 282
 FT DOMAIN 283 328
 FT DISULFID 127 140
 BY SIMILARITY.

FT DISULFID 134 149 BY SIMILARITY.
 FT DISULFID 151 162 BY SIMILARITY.
 FT DISULFID 168 177 BY SIMILARITY.
 FT DISULFID 173 186 BY SIMILARITY.
 FT DISULFID 188 201 BY SIMILARITY.
 FT DISULFID 207 217 BY SIMILARITY.
 FT DISULFID 213 226 BY SIMILARITY.
 FT DISULFID 228 241 BY SIMILARITY.
 FT DISULFID 247 258 BY SIMILARITY.
 FT DISULFID 254 267 BY SIMILARITY.
 FT DISULFID 269 281 BY SIMILARITY.
 FT DISULFID 287 300 BY SIMILARITY.
 FT DISULFID 294 309 BY SIMILARITY.
 FT DISULFID 315 327 BY SIMILARITY.
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 394 394 A -> T (IN REF. 1).
 FT CONFLICT 5 5 EMPDSDSH -> TOTAN (IN REF. 2).
 FT CONFLICT 44 51 AOHDPNCP -> VNTQPLPT (IN REF. 2).
 FT CONFLICT 103 111 C -> W (IN REF. 2).
 FT CONFLICT 294 294 RSV -> AER (IN REF. 2).
 FT CONFLICT 354 356 S -> R (IN REF. 3).
 FT CONFLICT 355 355
 SQ SEQUENCE 443 AA; 49391 MW; 9E9AC2393780D3B8 CRC64;
 Query Match 50.9%; Score 1289; DB 1; Length 443;
 Best Local Similarity 49.9%; Pred. No. 5.7e-91;
 Matches 226; Conservative 74; Mismatches 131; Indels 22; Gaps 5;
 QY 1 MPGIRILVTTLALCLPSPGNAQ-----AQTNGFDLDRSGGCLDIDECRTPEACRG 55
 DB 8 LFGSILLMALLLLLSASBPDSSEPDSTYECIDGYEMDSDSHCRVNECITPEACKG 67
 QY 56 DMACVNQNGCYICIPPTNVPYKGPYSPYSGPYAPAAPLAPNPTISPLICRF 115
 DB 68 EMKCIHNGYGLCLPRAAIVINDLHG-----EGP-PPVPPAQHFN-----PCPP 111
 QY 116 GROMESNOCVAVDECATDSHOCNPQICINTEGGYTCSTDGYWMLLEGGCLDIDECRYG 175
 DB 112 GYEPDDSCVVDYDECAQLHDCRPSQDCHNLPGSTQCTPDDGRTKIGPCVDDECRYR 171
 QY 176 YCOQLCANYPGSYCTCNPGFTLNEGRSCQVDNENATENPCVQTCVNTGSGFTCRDPG 235
 DB 172 YCOHRCVNLPGSFRCCGCEPFGQFGPNNRSCVDVNECDMGAPCQRCNNSGTFLCRHQG 231
 QY 236 YLEDEDVHGSCMDESESEFLCQHECVNQPGYFSCPPGYILLDNRSCODINEGHR 295
 DB 232 YELHRDGFSCSDIDECSSSYLCQYRCVNEPGRFSCHCPQGYQL-ATRCODIDECESG 290
 QY 296 NHTCNLOQTCVNLGGGFKCIDPIRCEPEYLRISDNRCMCFAENPGCRDOPFTILYRDMV 355
 DB 291 AHQCEAQTCAVPHFGYRCVDINRCYEPYIQVSENKCLCPASNPFLCREQSSIVHRYMTI 350
 QY 356 VSGRSVPADIFQMATTIRPGAYIYFIQKSGNREGRFYMRQTPISATVMTPRKPRE 415
 DB 351 TSERSVPADVFOIQATSVYPGAYNARQIRAGNSQGFYIQRINNVSAMVLAVPVGPRE 410
 QY 416 IQLDLEMTVNVINFRGSSVIRLRYVQYPR 448
 DB 411 YVLDLEMTVNSLMSYRASSVIRLTVFGAYTP 443
 RESULT 6
 FBLA_MOUSE STANDARD: PRT; 443 AA.
 AC 09WVY9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 (Fibulin-4) (FIBL-4) (Mutant p53 binding protein 1).
 GN EFEMP2 OR FBLN4 OR MBP1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Db 1 MPEGLRIITVLTALMLPHPNMAOQCCTNGFDLRQSGQCLDIDECRTIPEACRGDMKCV 60
QY 61 NONGGLICIPRTNPVYRGPSNPSTYSRGYPAAAPLSPNPPTISRLPLICRGYQMD 120
Db 61 NONGGLICIPRTNPVYRGPSNPSTYSRGYPAAAPLSPNPPTISRLPLICRGYQMD 120
QY 121 ESNOCVDEVCANDSHQCNPTQICINTEGGYCTCSDGTYWLEGGQCLDIDECRYGCOOL 180
Db 121 EGNQCVDEVCANDSHQCNPTQICINTEGGYCTCSDGTYWLEGGQCLDIDECRYGCOOL 180
QY 181 CANVPSTSYCTCNPFGTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPGELEE 240
Db 181 CANVPSTSYCTCNPFGTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPGELEE 240
QY 241 DGVCSDMECSSEFLCOHECVNQPGTYCSPPGYLLDNDNSCODINCEHNHCN 300
Db 241 DGVCSDMECSSEFLCOHECVNQPGTYCSPPGYLLDNDNSCODINCEHNHCN 300
QY 301 LOOTCYNLGGFKCIDPIRCEEPYLRISDNRCMCPAENPCGRDQPFYLLYRDMVYSGRS 360
Db 301 LOOTCYNLGGFKCIDPIRCEEPYLRISDNRCMCPAENPCGRDQPFYLLYRDMVYSGRS 360
QY 361 VPADIFOMATTPYRGAYTYFQIKSGNEGRFYMKGTPISATLVMPRIKGPRIQIDL 420
Db 361 VPADIFOMATTPYRGAYTYFQIKSGNEGRFYMKGTPISATLVMPRIKGPRIQIDL 420
QY 421 EMITVNTVINFGRSSVRLRIYVSQYPE 448
Db 421 EMITVNTVINFGRSSVRLRIYVSQYPE 448

```

RESULT 4
FBL4_CRIGR STANDARD: PRT: 443 AA.

```

AC 055058;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
DE (Fibulin-4) (FBLN-4) (H411 protein).
GN EFEMP2 OR FBLN4.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxId=10029;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
HE Helne H., Delude R.L., Monks B., Golanbeck D.T.;
SUBMITTED (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF046870; AAC03101.1;
CC HSSP: P00736; IAP0.
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000361; EGF-like.
CC Pfam: PF00008; EGF_4.
CC SMART: SM000179; EGF_CA; 4.
CC SMART: SM00001; EGF_Like; 2.
CC PROSITE: PS00010; ASX_HYDROXYL; 4.
CC PROSITE: PS00022; EGF_1; FALSE_NEG.
CC PROSITE: PS01186; EGF_2; 4.

```

```

DR PROSITE: PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 26 443
FT DOMAIN 36 81 EGF-Like 1, DIVERGENT.
FT DOMAIN 123 163 EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 164 202 EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 203 242 EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 243 282 EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 283 328 EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 127 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 151 162 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 173 186 BY SIMILARITY.
FT DISULFID 188 201 BY SIMILARITY.
FT DISULFID 207 217 BY SIMILARITY.
FT DISULFID 213 226 BY SIMILARITY.
FT DISULFID 228 241 BY SIMILARITY.
FT DISULFID 247 258 BY SIMILARITY.
FT DISULFID 254 267 BY SIMILARITY.
FT DISULFID 269 281 BY SIMILARITY.
FT DISULFID 287 300 BY SIMILARITY.
FT DISULFID 294 309 BY SIMILARITY.
FT DISULFID 315 327 BY SIMILARITY.
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 443 AA; 49432 MW; 0BCEFD7323D9E5F CR64;

```

Query Match 50.9%; Score 1290; DB 1; Length 443;
Best Local Similarity 50.1%; Pred. No. 4.8e-91;
Matches 227; Conservative 71; Mismatches 133; Indels 22; Gaps 5;

```

QY 1 MPEGLRIITVLTALMLPHPNMAOQCCTNGFDLRQSGQCLDIDECRTIPEACRG 55
Db 8 IIPSLILMALILILGLAASPDQSEPDSTCTGCEYEMDADSOHCRDVNECLTIEACRG 67
QY 56 DMNCVNOGGYLCIPRTNPVYRGPSNPSTYSRGYPAAAPLSPNPPTISRLPLICRF 115
Db 68 EMKICINHYGTYLCLPRSAVYNDLHG-----EGP-PPVPPAQHPN-----PCPP 111
QY 116 GYOMDESNOGVDECATDHSQCNPQICINTEGGYCTCSDGTYWLEGGQCLDIDECRYG 175
Db 112 GYEPDEGSCVDVDECAQALHDCRPSQDCHMLPGSYCTCPDGYRKVGPEVDIECRVR 171
QY 176 YCOQLCANVPSTSYCTCNPFGTLNEDGRSCQDVNCEATENPCVQTCVNTYGSFTICRDPG 235
Db 172 YCOHRCVNLGFSFRCQCEPFGOLGPNNRSCYDVNECDMGAPCEORCFNSYGTFLCRNOG 231
QY 236 YELEEDGVHCSDMDECSSEFLCOHECVNQPGTYCSPPGYLLDNDNSCODINCEHR 295
Db 232 YEHLRDFSCSDIDECYSSTLCQRYCNEVNEGRSCHCPQGYQL-ATRLCODIDECETG 290
QY 296 NHTCNLOOTCYNLGGFKCIDPIRCEEPYLRISDNRCMCPAENPCGRDQPFYLLYRDMV 355
Db 291 AHQCSAQTCVNFHGRVDTNRCVEPYGVSDMRCPSPNPLCREQPSIVHRYSI 350
QY 356 VSGRSVPADIFOMATTPYRGAYTYFQIKSGNEGRFYMKGTPISATLVMPRIKGPRI 415
Db 351 TSEKSVPADVFOIQTATSVYRGAYNAFOIRAGNTGDFYROIINVSAMLVLRAPVTPGRE 410
QY 416 IOLDLEMITVNTVINFGRSSVRLRIYVSQYPE 448
Db 411 YVLDLEMITVNTVINFGRSSVRLRIYVSQYPE 443

```

RESULT 5
FBL4_HUMAN STANDARD: PRT: 443 AA.
AC 095967; 075967;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT SITE 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 238 238 L -> P (IN REF. 2).
SQ SEQUENCE 448 AA: 50160 MM: E6BC68F7B14B714 CRC64:

Query Match 95.08; Score 2406; DB 1; Length 448;
Best Local Similarity 94.28; Pred. No. 7.8e-176;
Matches 422; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

OY 1 MFGIRLITVLTIALLPSPGNNAOCTNFDLDROSGOCLDIDECTIPACRGDMVCV 60
DB 1 MFGRLITVLTIALLPSPGNNAOCTNFDLDROSGOCLDIDECTIPACRGDMVCV 60
OY 61 NONGCYLCPRTNPVYRGYPNPSYSGPYPAAAPPLSAPNYPPTISRLPCLICRFGYOMD 120
DB 61 NONGCYLCPRTNPVYRGYPNPSYSGPYPAAAPPLSAPNYPPTISRLPCLICRFGYOMD 120
OY 121 ESNOCVVDDECATDHSQCNPOTICINTEGSGYSCDIDGWLLEGOCLDIDECRYGCOOL 180
DB 121 ESNOCVVDDECATDHSQCNPOTICINTEGSGYSCDIDGWLLEGOCLDIDECRYGCOOL 180
OY 121 EGNOCVVDDECATDHSQCNPOTICINTEGSGYSCDIDGWLLEGOCLDIDECRYGCOOL 180
DB 121 EGNOCVVDDECATDHSQCNPOTICINTEGSGYSCDIDGWLLEGOCLDIDECRYGCOOL 180
OY 181 CANVGSYSCTNPGFTLNEDEGRSCODVNCATENPCVOCVNTYSGSFICRCDGYELE 240
DB 181 CANVGSYSCTNPGFTLNEDEGRSCODVNCATENPCVOCVNTYSGSFICRCDGYELE 240
OY 181 CANVGSYSCTNPGFTLNEDEGRSCODVNCATENPCVOCVNTYSGSFICRCDGYELE 240
DB 181 CANVGSYSCTNPGFTLNEDEGRSCODVNCATENPCVOCVNTYSGSFICRCDGYELE 240
OY 241 DGVHCSMDMECSFSEFLCOHECVNOGPGYFCSPGPIILDDNRSODINECHRNHTCN 300
DB 241 DGVHCSMDMECSFSEFLCOHECVNOGPGYFCSPGPIILDDNRSODINECHRNHTCN 300
OY 241 DGHCSMDMECSFSEFLCOHECVNOGPGYFCSPGPIILDDNRSODINECHRNHTCT 300
DB 241 DGHCSMDMECSFSEFLCOHECVNOGPGYFCSPGPIILDDNRSODINECHRNHTCT 300
OY 301 LOQTCYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDPFTILYRDMVSGRS 360
DB 301 LOQTCYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDPFTILYRDMVSGRS 360
OY 301 PLOTCTYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDPFTILYRDMVSGRS 360
DB 301 PLOTCTYNLQGGFKCIDPIRCEPEYLRISDNRCMCPAENPCRCRDPFTILYRDMVSGRS 360
OY 361 VPADIFQWQATRRYPGAYIYIFQIKSGNEGREFYMRQGPISATLVMTRPKIGPRDIQDL 420
DB 361 VPADIFQWQATRRYPGAYIYIFQIKSGNEGREFYMRQGPISATLVMTRPKIGPRDIQDL 420
OY 421 EMITVNTVYINFRGSSVYLRIRIYSOXP 448
DB 421 EMITVNTVYINFRGSSVYLRIRIYSOXP 448
OY 421 EMITVNTVYINFRGSSVYLRIRIYSOXP 448
DB 421 EMITVNTVYINFRGSSVYLRIRIYSOXP 448

RESULT 3
FBL5_MOUSE STANDARD: PRT; 448 AA.
AC Q9WYH9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Ftblin-5 precursor (Ftbl-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance).
GN FBLN5 OR DANCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE, a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF112151; AAD41767.1; -.
CC HSP: P00736; IAP0.
CC MGD: MGI:1346091; Fbln5.
CC InterPro: IPR000152; Asx_hydroxyl.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR001881; EGF_Ca.
CC Pfam: PF00008; EGF_4.
CC SMART: SM000179; EGF_CA: 4.
CC SMART: SM00001; EGF_Like: 2.
CC PROSITE: PS00010; ASX_HYDROXYL; 4.
CC PROSITE: PS00022; EGF_1; FALSE_NEG.
CC PROSITE: PS01186; EGF_2; 4.
CC PROSITE: PS01187; EGF_CA: 6.
CC Repeat: EGF-like domain; Calcium-binding; Glycoprotein; Signal.
CC SIGNAL 1 23
CC CHAIN 24 448
CC DOMAIN 24 69 FIBULIN-5.
CC DOMAIN 127 167 EGF-LIKE 1, DIVERGENT.
CC DOMAIN 168 206 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
CC DOMAIN 207 246 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
CC DOMAIN 247 287 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
CC DOMAIN 288 333 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
CC SITE 54 56 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
CC DISULFID 131 144 CELL ATTACHMENT SITE (POTENTIAL).
CC DISULFID 138 153 BY SIMILARITY.
CC DISULFID 155 166 BY SIMILARITY.
CC DISULFID 172 181 BY SIMILARITY.
CC DISULFID 177 190 BY SIMILARITY.
CC DISULFID 192 205 BY SIMILARITY.
CC DISULFID 211 221 BY SIMILARITY.
CC DISULFID 217 230 BY SIMILARITY.
CC DISULFID 222 245 BY SIMILARITY.
CC DISULFID 232 245 BY SIMILARITY.
CC DISULFID 251 262 BY SIMILARITY.
CC DISULFID 258 271 BY SIMILARITY.
CC DISULFID 273 286 BY SIMILARITY.
CC DISULFID 292 305 BY SIMILARITY.
CC DISULFID 299 314 BY SIMILARITY.
CC DISULFID 320 332 BY SIMILARITY.
CC CARBOHYD 283 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 448 AA: 50193 MM: F15CC70CCFBDC97 CRC64:

Query Match 94.98; Score 2405; DB 1; Length 448;
Best Local Similarity 94.28; Pred. No. 9.3e-176;
Matches 422; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

OY 1 MFGIRLITVLTIALLPSPGNNAOCTNFDLDROSGOCLDIDECTIPACRGDMVCV 60

```

```

DR Genew; HGNC:3602; FBLSN.
DR MIM; 604580; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_Like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
FT CHAIN 1 23 POTENTIAL.
FT SIGNAL 1 23 POTENTIAL.
FT DOMAIN 24 448 EGF-Like 1, DIVERGENT.
FT DOMAIN 127 167 EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 69 70 IP -> HS (IN REF. 3).
FT CONFLICT 147 148 TE -> MS (IN REF. 3).
SQ SEQUENCE 448 AA; 50180 MW; 19FCA51FDA328003 CRC64;

```

```

Query Match 100.0%; Score 2533; DB 1; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.8e-185;
Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 361 VPADIFPMQATRRPGAYYIFQIKSGNCRGFYMQTGPISATLVNTRPIKGPRTQLDL 420
QY 421 EMITVNTVYINFRGSSVIRLRITYVSQYP 448
DB 421 EMITVNTVYINFRGSSVIRLRITYVSQYP 448

RESULT 2
FBLS_RAT STANDARD: PRT: 448 AA.
AC Q9WVH8; Q9R284;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibulin-5 precursor (FBLN-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance) (Embryonic vascular EGF repeat-containing
DE protein) (EVEC).
GN FBLN5 OR DANCE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE, a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99278197; PubMed=10347091;
RA Kowal R.C., Richardson J.A., Milano J.M., Olson E.N.;
RT "Evec, a novel epidermal growth factor-like repeat-containing protein
RT upregulated in embryonic and diseased adult vasculature.";
RL Circ. Res. 84:1166-1176(1999).
CC 1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC 1- SUBCELLULAR LOCATION: Secreted.
CC 1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF112153; AAD41769.1; -.
DR EMBL; AF137350; AAD25101.1; -.
DR HSSP; P00736; IAPQ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-Like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_Like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 6.
KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 448 FIBULIN-5.
FT DOMAIN 127 167 EGF-Like 1, DIVERGENT.
FT DOMAIN 168 206 EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-Like 4, CALCIUM-BINDING (POTENTIAL).

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:40:07 ; Search time 11.3157 seconds

(without alignments)
1642.086 Million cell updates/sec

Title: US-09-674-379A-13
Perfect score: 2533
Sequence: 1 MPGIKRLITVITLALCLPSP.....INFRGSSVIRLRIVSQYPF 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2533	100.0	448	1 FBL5_HUMAN	Q9ubx5 homo sapien
2	2406	95.0	448	1 FBL5_RAT	Q9wv18 rattus norv
3	2405	94.9	448	1 FBL5_MOUSE	Q9wv19 mus musculu
4	1290	50.9	443	1 FBL4_CRIGR	O55058 cricetulus
5	1289	50.9	443	1 FBL4_HUMAN	O95967 homo sapien
6	1282	50.6	443	1 FBL4_MOUSE	O9wv19 mus musculu
7	1109.5	43.8	493	1 FBL3_RAT	O35568 rattus norv
8	1101	43.5	493	1 FBL3_HUMAN	O12805 homo sapien
9	778.5	30.7	684	1 FBL1_CHICK	O73775 gallus gall
10	733.5	29.0	11221	1 FBL2_MOUSE	P37889 mus musculu
11	709.5	28.0	1184	1 FBL2_HUMAN	P38095 homo sapien
12	703.5	27.8	703	1 FBL1_MOUSE	O08879 mus musculu
13	698.5	27.5	703	1 FBL1_HUMAN	P23142 homo sapien
14	599.5	23.7	712	1 FBL1_CAEEL	O77469 caenorhadi
15	544	21.5	1394	1 LTRBS_HUMAN	P22064 homo sapien
16	544	21.5	1595	1 LTRBL_HUMAN	O14766 homo sapien
17	533.5	21.0	2871	1 FBN1_MOUSE	P35555 homo sapien
18	531.5	21.0	2871	1 FBN1_BOVIN	P98133 bos taurus
19	525.5	20.7	2871	1 FBN1_PIG	O91v16 sus scrofa
20	523	20.6	1712	1 LTRBL_RAT	O00918 rattus norv
21	518.5	20.5	2871	1 FBN1_MOUSE	O61554 mus musculu
22	518.5	20.5	2907	1 FBN2_MOUSE	O61555 mus musculu
23	517.5	20.4	2911	1 FBN2_HUMAN	P35556 mus musculu
24	475	18.8	956	1 MTN2_HUMAN	O00339 homo sapien
25	452	17.8	956	1 MTN2_MOUSE	O08766 mus musculu
26	421	16.6	931	1 EMRI_MOUSE	O61349 mus musculu
27	392	15.5	886	1 EMRI_HUMAN	O14246 homo sapien
28	379	15.0	2470	1 NTC2_MOUSE	O35516 mus musculu
29	377	14.9	2471	1 NTC2_HUMAN	O04721 homo sapien
30	375	14.8	810	1 NEI2_HUMAN	O92832 homo sapien
31	372.5	14.7	816	1 NEI2_MOUSE	O61220 mus musculu
32	371.5	14.7	2471	1 NTC2_RAT	O9qw30 rattus norv
33	366	14.4	816	1 NEI2_HUMAN	O99435 homo sapien

34	364	14.4	810	1 NEI1_RAT	O62919 rattus norv
35	360.5	14.2	816	1 NEI1_CHICK	O90827 gallus gall
36	359.5	14.2	1964	1 NTC4_MOUSE	P31695 mus musculu
37	359	14.2	816	1 NEI2_RAT	O62918 rattus norv
38	356	14.1	2703	1 NTC1_DROME	P07207 drosophila
39	355.5	14.0	2437	1 NTC1_BRARE	P46530 brachydanio
40	350.5	13.8	652	1 CD93_HUMAN	O9npv3 homo sapien
41	347.5	13.7	2556	1 NTC1_HUMAN	P46531 mus sapien
42	347	13.7	644	1 CD93_MOUSE	O89103 mus musculu
43	346	13.7	1064	1 FBL1_STRPU	P10079 strongyloce
44	346	13.7	2524	1 NTC1_XENLA	P21783 xenopus lae
45	344	13.6	2321	1 NTC3_HUMAN	O9um47 homo sapien

ALIGNMENTS

RESULT 1
FBL5_HUMAN STANDARD; PRT; 448 AA.
ID FBL5_HUMAN
AC Q9UBX5; 075966;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fibulin-5 precursor (FBL-5) (developmental arteries and neural crest
EGF-like protein) (Dance) (Urine p50 protein) (UP50).
GN FBLN5 OR DANCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Kostka G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428823;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE", a novel secreted RGD protein expressed in developing,
atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=urine;
RA Zemel R., Sholto O., Shaul Y.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
DEVELOPMENT AND REMODELING.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND
COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.
CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL
BLOOD LEUCOCYTES.
CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AJ133490; CAB38568.1; -
CC EMBL: AF112152; AAD41768.1; -
CC EMBL: AF093118; AAC62107.1; -
CC HSP; P00736; IADP.

```
OY 217 -----CVQ-----TCVNTYGFICRCDPGYELEDGVHC 245
DB 1392 YRCLCYDGFMAEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDMGYSKKRGKTC 1451
OY 246 SDMECSSEFLC-QHE-CVNQGTFCSPGPIILLDNRSCODINECEHRNHTCNLQ 303
DB 1452 TDINCEIGAHNCGRHACTNAGSFKCSGPMI--GDGIKCTDLDECSTNGTMCQHA 1509
OY 304 TCYNLQGGFKCIDPIRCEPYL-----RISDRCMCPAENPGCRDOP 345
DB 1510 DCKNTMGSYRCL-----CKEGYTGDFCTDLDCESENILNC--GNGOCLNAP 1555
```

RESULT 15

```
A55567
fibrillin I - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002
C:Accession: A55567
R:Listra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
Title: Sequence of the coding region of the bovine fibrillin CDNA and localization to
Reference number: A55567; MUID:95137597; PMID:7835900
Accession: A55567
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <Till>
C:Cross-references: GB:L28748; NID:9508427; PIDN:AAA74122.1; PID:9508428
C:Superfamily: fibrillin 1; EGF homology
F:1201-1236/Domain: EGF homology <EGF>
```

```
Query Match 21.0%; Score 531.5; DB 2; Length 2871;
Best Local Similarity 29.6%; Pred. No. 4,4e-27;
Matches 122; Conservative 48; Mismatches 129; Indels 113; Gaps 16;

OY 16 CLPSPGNAQACTNGFDLDQSGQCLDIDECRTIPACRGDMCVNONGYLCPRTNPV 75
DB 1044 CRNTIGSFRCRCDGFDLSEERNCTDIDECRISPLIC--GRGQCVNTPGDFEC--KCDEG 1100
OY 76 YRGPSNPSYSTRSGPYPAAPPLSAPNPTISR---PLICRGYOMDES----- 122
DB 1101 YESGF-----MMKNCMDIDECQRPDLRCRGVCIATGEGSYRCGCP 1142
OY 123 -----NOCVDECATDSHCNPTQICINTEGGYCSCTDGYWLLSQ--CLDIDEC 172
DB 1143 GHQIAPNISACIDINECELSAHLIC--PHGRCVNLIGKYQCAQCNPGYHSTPRLRCVDIDEC 1201
OY 173 RY--GYCOQLCANVPQSYSCNPGFTLNEDGRSCODVNECATENP----- 216
DB 1202 SIMNGCETFTNSGSEYSCQPGFALMPDQSRCTDIDEC--EDNPNICDGGQCTNIPGE 1260
OY 217 -----CVQ-----TCVNTYGFICRCDPGYELEDGVHC 245
DB 1261 YRCLCYDGFMAEDMKTCVDVNECDLNPNICLSGTCENTKGSFICHCDMGYSKKRGKTC 1320
OY 246 SDMECSSEFLCQHE--CVNQGTFCSPGPIILLDNRSCODINECEHRNHTCNLQ 303
DB 1321 TDINCEIGAHNCGRHACTNAGSFKCSGPMI--GDGIKCTDLDECSTNGTMCQHA 1378
OY 304 TCYNLQGGFKCIDPIRCEPYL-----RISDRCMCPAENPGCRDOP 345
DB 1379 DCKNTMGSYRCL-----CKEGYTGDFCTDLDCESENILNC--GNGOCLNAP 1424
```

Search completed: July 3, 2003, 18:22:41
Job time : 20.0023 secs

transforming growth factor beta-1-binding protein - human

C:Species: Homo sapiens (man)
C:Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 11-Jan-2000
C:Accession: A35626
R:Kanaki, T.; Olofsson, A.; Moren, A.; Wernstedt, C.; Hellman, U.; Miyazono, K.; Claess
Cell 61, 1051-1061, 1990
A:Title: TGF-beta1 binding protein: a component of the large latent complex of TGF-beta1
A:Reference number: A35626; MUID:90275601; PMID:2350783
A:Accession: A35626
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1394 <KAN>
A:Cross-references: GB:M34057; NID:9339547; PIDN:AAA61160.1; PID:9339548
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing
E:750-791/Domain: EGF homology <EGF>

Query Match 21.5% Score 544; DB 2; Length 1394;

Best Local Similarity 29.4%; Pred. No. 3.5e-28;

Matches 128; Conservative 52; Mismatches 134; Indels 122; Gaps 17;

16 CLPSPGNAOCTNGEDLDROSGCCLDIDECRTPEACRGDMCVNONGYLCPRTNPV 75
119 MDESNOCVDDECATSDHOCNPQICINTEGCTC-SCDGVWLLGOCCLDIDEC-RGY 176
663 MNGRCEDIDECLNPS-TC-PDEQCVNSGSGVCPCTEEFRGMNGCLDVEDELENNV 720
177 CQO-LCANPGSGSCINPFTLNEDGRSCQDVNECATENPCVQ----- 219
721 CANGCDSNLESGSYSCGKGYTRPHKHCRIIDECQOQNCVNGQCKNTEGSEFCTCGQ 780
220 -----TCVNTGSGFICRCDPGYLEEDGVHCSMDDECSF 253
781 GYOLSAKNOCEIDECQNHLCAGHCCRTGSEFCVCCDGYGASGIGDCEIDNECLE 840
254 SEFLCQH-ECVNGPGTYFCSPGXYLLDDNRSCQDINECRHNTCMACQTCYNLAGGF 312
841 DRSVCQRGDCINAGSYDCTCPDGF-QLDNNKTCQDINECHPG-ICGPGEGCLNTEGSEF 898
313 KCL-----DPRICEPYLRIS-----DN-----RMC-----PAENPCR 342
899 HCVCCQGFISAGRTCEDIDECVNNTVCDSHGFCDDNTAGSFRCICYGFOAPDGGCCV 958
343 DQPFILYRDMDVNSG 358
959 D-----VNECELLSG 968

RESULT 14

A47221 fibrillin 1 precursor - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 02-Jun-1995 #sequence_revision 25-Apr-1997 #text_change 02-Aug-2002
C:Accession: A47221; S17064; S17064; S17064; S62111; A34198
R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
Genomics 17, 476-484, 1993
A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
A:Reference number: A47221; MUID:94010947; PMID:7691719
A:Accession: A47221
A:Molecule type: mRNA
A:Residues: 1-337, 'T', 339-1029 <COR>
A:Cross-references: GB:X63556
R:Perle, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillman, T.; Bonad
Hum. Mol. Genet. 2, 961-968, 1993
A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
A:Reference number: I54355; MUID:93372860; PMID:8364578
A:Accession: I54355

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 132-3002 <PER>
A:Cross-references: GB:L13923; NID:9306745; PIDN:AAB02036.1; PID:9306746
R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
Nature 352, 334-337, 1991
A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
A:Reference number: S17064; MUID:91304568; PMID:1852207
A:Accession: S17064
A:Molecule type: mRNA
A:Residues: 1030-3002 <MAS>
A:Cross-references: EMBL:X63556
R:Dietz, H.C.; Valle, D.; Francmanno, C.A.; Kendzior, R.J.
Science 259, 680-683, 1993
A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
A:Reference number: I59574; MUID:93157831; PMID:8430317
A:Accession: I59574
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 2217-2288, 'T', 2290-2325 <RES>
A:Cross-references: GB:S54426; NID:9264860; PIDN:AAB25244.1; PID:9264861
R:Lee, B.; Godfrey, M.; Vitale, E.; Hort, H.; Mattei, M.G.; Sarrafz, M.; Tsipouras,
Nature 352, 330-334, 1991
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
A:Reference number: S17062; MUID:91304567; PMID:1852206
A:Accession: S17062
A:Molecule type: mRNA
A:Residues: VLVTVVFIFSYNKM, 944-1444 <LEE1>
A:Cross-references: EMBL:X62008; NID:931398; PIDN:CAB56534.1; PID:95924015
A:Accession: S62111
A:Molecule type: protein
A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>
R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
J. Biol. Chem. 264, 21381-21385, 1989
A:Title: Connective tissue microfibrils. Isolation and characterization of three larg
A:Reference number: A34198; MUID:90078246; PMID:2512293
A:Accession: A34198
A:Molecule type: protein
A:Residues: 565-575, 1890-1892, 'I', 1894-1900 <MAD>
C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
C:Genetics:
A:Gene: GDB:FBNI
A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
A:Map position: 15q21.1-15q21.1
A:Introns: 2236/1; 2258/1; 2297/1
C:Superfamily: fibrillin 1; EGF homology
C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein
F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predict
F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MATC>
F:1332-1367/Domain: EGF homology <EGF>
F:1457-1492/Domain: EGF homology <EGF2>
F:2262-2295/Domain: EGF homology <EGF1>

Query Match 21.0% Score 532.5; DB 2; Length 3002;

Best Local Similarity 30.1%; Pred. No. 3.9e-27;

Matches 124; Conservative 50; Mismatches 125; Indels 113; Gaps 18;

16 CLPSPGNAOCTNGEDLDROSGCCLDIDECRTPEACRGDMCVNONGYLCPRTNPV 75
1175 CRTTISFCRCDSGFALDSEERNCTDIDECRISPDIC-GRGQCVNPPDDECC-KCDEG 1231
76 YRGPNPNPSTPSGYPAPAPLAPNPTISR-----PLICRF----- 115
1232 YESGF-----MMKNCMDIDECQDPLCRCGVCNTEGSTRCECP 1273
116 GYOMDES-NOCVDDECATSDHOCNPQICINTEGCTCSDGVWLLGOCCLDIDEC 172
1274 GQDLSNINISACIDINECELSAHLIC-PNGRKNVNLGKQACNPNYHSPDLFCVDIDEC 1332
173 RV-GYQOOLCANVPGSYSCINPFTLNEDGRSCQDVNECATENP----- 216
1333 SIMNGCECTFCTNSESSEYSCQPGFALMDORSCRTDIDEC-EDNPNICDGGCTNIPGE 1391

RESULT 11

T43210

fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character

A:Reference number: Z23337

A:Accession: T43210

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 23.5%; Score 596; DB 2; Length 589;

Best Local Similarity 27.3%; Pred. No. 6,3e-32;

Matches 159; Conservative 53; Mismatches 163; Indels 208; Gaps 21;

15 ICLPSPG-NAQAQCTNGFDLDRSGGCLD-IDECRTIPACRGDMCMVONGGYLCIPRT 72

66 LCHDRGKGVKSCRSRSGFDLAPDGMACVDRHIDECATLMDCLLESQRCLTPGSEFKCI 122

73 NPVYRGPNPSTPYSGPYPAAPPLASAPNPTISPLICRPGYQMD-ESNOCVDVDEC 131

123 -----FTLSCGTGYAMSETERNCVDVDEC 146

132 ATDSHOCNPTQICINTEGGYTC-----SCTDGYWLEGGCLDID 170

147 NLGSHCCGLLYOCRNTOGSGYRCDAKKGDELQNPMTGECTSTCPNRYRKNGMCNID 206

171 ECRYGT-----CQO---LCA----- 182

207 ECVTGHCNCGAGECVNTPGSEFCQKGNLCAHGEVNGATGFCEDVNECQGVGSMECT 266

183 NPGSYSCNPNPFTLNE----- 200

267 NLPRTKCKCGPGEYENDAKKRCEDVDECITKFAHVCDDLSEACINTIGSEFCKCKPGFOL 326

201 --DGRSCQDVNECATE-NPCVQTCVNTYGSFICRCDPGYLEEDGVHSCDMECSF---- 253

327 ASDGRCEVNECTTGIAACEQKCVNTPGSGYQCICDGFALGPGTGKEDIDECIMAGS 386

254 SEFLCOHECVNPGYFCSCPGYIILADNRSCODINECEHNRHTCNLQOTCYNLQGGFK 313

387 GNDLCMGCCINTKGYLCOCPPGKTIQPDGRTCVDVDECA-MGECAGSDKVCVNTLGSFK 445

314 CIDPIRCEPPLRLISNR-----C-MCPAENPGC-RDQPTILYRMDVVSGRSV- 361

446 C-HSIDCPNTYIHDSLNKNQIADGYSCIKVSTEDTCLGNHTREVLQFRAVPSLKTII 504

362 -PADI-----FOMQATRYPGAYYIFQIKSGNEGRFYMROTPISATLVMPRI 410

505 SPIEVSRIYTHMGVPSVDNLDYVQGRHRIYQERNG-----IYQLVAPRI 551

411 KGPRIQLDLEMTVNTVINFGRSSVIR-----LRIYSQYRF 448

552 SGP-----YVERIKVINHTKSRFGVILAFNEAIEISVSKYF 589

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

A:Reference number: Z19616
 A:Accession: T22793
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-798 <WIL>
 A:Cross-references: EMBL:Z68749; PIDN:CAA92962.1; GSPDB:GN00022; CESP:F56H11.1
 A:Experimental source: clone F56H11
 R:Lloyd, C.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: Z19897
 A:Accession: T24489
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-798 <WIL>
 A:Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GN00022; CESP:F56H11.1
 A:Experimental source: clone T05A1
 C:Genetics:
 A:Gene: CESP:F56H11.1
 A:Map position: 4
 A:Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3; 498/3;
 C:Superfamily: fibulin-1; EGF homology

Query Match 22.8%; Score 577.5; DB 2; Length 798;
 Best Local Similarity 26.7%; Pred. No. 1.4e-30;
 Matches 162; Conservative 50; Mismatches 180; Indels 215; Gaps 23;

15 ICLPSPG-NAQAQCTNGFDLDRSGGCLDIDECRTIPACRGDMCMVONGGYLC----- 68
 167 LCHDRGKGVKSCRSRSGFDLAPDGMACVDRHIDECATLMDCLLESQRCLTPGSEFKCI 226
 69 -IP--RTNPVYRG--YSNPYSTPYSGPYPAAP----- 98
 227 RLVPYHRRANRIGNAPRRMRDPPYSR-AGEYREASQANTERGCWGLFQGHCVYDIE 284
 99 -----LSAPNPTISPLICRPGYQMD-ESNOCV----- 126
 285 CATLMDCLLESQRCLTPGSEFKCIITLSCGTGYAMSETERNCFLIILNTFNCKYFV 344
 127 -DVDECATDSHOCNPTQICINTEGGYTC-----SCTDGYWLEGGCLDID 154
 345 EDVDECMGSHDGPRLVOCRNTOGSGYRCDAKKGDELQNPMTGEXIDECVYGHNGAGE 404
 155 -----CTDGYWL--LEGGCLDIDECRYGCOQ--CANVPGSYSCNMP 194
 405 ECYNTPGSEFCQKGNLCAHGEVNGATGFCEDVNECQGVGSMECTNLPETYKCKGCP 464
 195 GFTLNE----- 210
 465 GYEFNDKKRCEDVDECITKFAHVCDDLSEACINTIGSEFCKCKPGFOLASDRRCEDVNE 524
 211 CATE-NPCVQTCVNTYGSFICRCDPGYLEEDGVHSCDMECSF----SEFLCOHECVNQ 265
 525 CTGIAACCEKCVNTPGSGYQCICDGFALGPGTGKEDIDECIMAGSNDLGMGCINT 584
 266 PGYFCSPGPGYIILADNRSCODINECEHNRHTCNLQOTCYNLQGGFKCIDPIRCEPYL 325
 585 KGSYLCQCPGKTIQPDGRTCVDVDECA-MGECAGSDKVCVNTLGSFKC-HSIDCPNTYI 642
 326 RISDNR-----CMCPAENPGC-RDQPTILYRMDVVSGRSV----- 362
 643 HDSLNNKRCNROPSACGILPEE--CSKVPLELTJYFISL--ARAVPISHRAPIYLFKVS 697
 363 ----ADI-----FOMQATRYPGAYYIFQIKSGNEGRFYMROTP--ISATLVMPRIKGP 413
 698 APNHADLEVFEIQTKTIYGAIPVLAIRAN-----FLQNGEKRNASVYILRSLDGP 752
 414 REIQDL 420
 753 QTVKQL 759

RESULT 13

A35626

hypothetical protein F56H11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002

C:Accession: T22793; T24489

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Title: Fibulin, a novel protein that interacts with the fibronectin receptor beta-subunit
 A:Accession: A32826; MUID:89354537; PMID:2527614
 A:Residues: 30-35, 'SX', 38-40, 'SH', 43-44 <AR3>
 A:Molecule type: protein
 C:Genetics:
 A:Gene: GDB:FBLN1; FBLN
 A:Cross-references: GDB:278285; OMIM:135820
 A:Map position: 22q13.3-22q13.3
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; glycoprotein
 F:1-29/Domain: signal sequence #status predicted <Sig>
 F:30-683/Product: fibulin 1 splice form C #status predicted <Mat>
 F:180-214/Domain: EGF homology <EGF>
 F:485-523/Domain: EGF homology <EGF>
 F:59,535/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.1%; Score 737; DB 2; Length 683;
 Best Local Similarity 35.4%; Pred. No. 4,1e-41;
 Matches 157; Conservative 67; Mismatches 154; Indels 66; Gaps 17;

QY 16 CLSPGNAQACTNGFLLDROSGCLDIDECRTIPEACRGDMCVNNGYLCTPRTNPV 75
 DB 288 CRP-----KLCCKSGFLTD-ALGNCIDINECLISAPCPICHTCINTEGSYTC----- 334
 QY 76 YRGPYNPTSPYSGPYPAAPPLSAPNYPITISRLICRFGYOM-DESNOCVDVDECATD 134
 DB 335 -----QKNVPR-----CGRTYHLNEBSTRCYVDVDECAP 363
 QY 135 SHOCNPTQICINTEGGYTCSTDGWY--LEGOCLDIDEC-RY-GYCOOLCANVGSYS 189
 DB 364 AEECGKGRHCVNPSGSPRCEKCTGYFDGISRMCCVDVDECCORYPRLCGHNCENTLGSTL 423
 QY 190 CTGNPGFTLNEDRSQCDVNECATENPCVQTCVNTYGSFICRCDPGYLEE-DEYHCSDM 248
 DB 424 CSOSVGRSLSDVDRSCDINECSS-SPSCQECANVGSYOCYCRGQYLDSDVGTCEDI 482
 QY 249 DECSF--SEFLCOHECVNOPTGYFCSCP-PGYTILDDNRSCODINECHRNHTCNLOOTC 305
 DB 483 DECALPTGHCIGYRCINIGSFQCSPPSSGYRLAPNRKQDDIDECYTGTHNCSINEEC 542
 QY 306 YNCGGFKCIDPTICEEPYRLISDNC-MCPA-ENPCGRDPFTILYRDMVYSGRSVPA 363
 DB 543 FNIQGFRCFL-AEECEPNYRRSATRCERLPCHENRECSKLPRLRTYVHLSFPTIQAPA 601
 QY 364 DIFOMATRYPGAYYIFQIKSGNEGEFYMROGPISTATVMTPRPKGPREIQIDLEMI 423
 DB 602 VYFRMGPSSAVPEDSQMLATGGNEGEFTTRKVPSPSGVALTKPVEPRDL-----LL 656
 QY 424 TVNTVINFNG--SSVIRLRIVYS 444
 DB 657 TVKMDLSRHGTVSSFAKLFIFVS 680

RESULT 5
 A:Accession: A49457
 A:Title: fibulin-2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Aug-2002
 C:Accession: A49457; S74095
 R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.
 J. Cell Biol. 123, 1269-1277, 1993
 A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with
 A:Reference number: A49457; MUID:94064787; PMID:8245130
 A:Accession: A49457
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met
 A:Reference number: S74094; MUID:96439073; PMID:8841408

A:Accession: S74095
 A:Molecule type: protein
 A:Residues: 236-238, 'X', 240-247, 260-275, 336-344, 'L', 346-361, 405-426, 566-568, 'EW', 569-
 C:Superfamily: fibulin-2; EGF homology
 C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer
 F:942-978/Domain: EGF homology <EGF>

Query Match 29.0%; Score 733.5; DB 2; Length 1221;
 Best Local Similarity 35.3%; Pred. No. 1.2e-40;
 Matches 146; Conservative 60; Mismatches 149; Indels 59; Gaps 11;

QY 23 AQACQTNFGLDROSGCLDIDECRTIPEACRGDMCVNNGYLCTPRTNPVYRGSYN 82
 DB 829 ARQRCMDGF-LQDEGNCVDINECTSLIEPCRSRSGFSINTGYSYTC----- 873
 QY 83 PYSTPYSGPYPAAPPLSAPNYPITISRLICRFGYOM-DESNOCVDVDECATDHSQNP 141
 DB 874 -----QKNVPR-----CGRTYHLNEBSTRCYVDVDECAP 363
 QY 142 QICINTEGGYTCSTDGWY--LEGOCLDIDECRYG--YCOOLCANVGSYSCTCNP 196
 DB 910 QLCYNLPGSYRCDCKPGFORDAFGRTCIDVNECVWSPGRLCQHNCENTPGYRSCAAGF 969
 QY 197 TLNEDRSQCDVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMDSCS-FSE 255
 DB 970 LNAADGRHCDVNECTERR-CSQECANVGSYOCYCRGQYLAEDGHTCTDIDECQAG 1028
 QY 256 FLCOHCVNOPTGYFCSCP-PGYTILDDNRSCODINECHRNHTCNLOOTCYNLOGGRC 314
 DB 1029 ILCTFPICVNVPSGYOCACPPQGTMTMANGSKDLDCALCTHNCSEMETHNIOGSR 1088
 QY 315 IDPRCEEPYRLISDNCMPAENPCGRD-----QPTILYRDMVYSGRSVPA 368
 DB 1089 L-RFDCPPNVYRVSQRC-----ERTTCQDITECATSPARITVYQNLNFTQGLVLAHPIRI 1143
 QY 369 QATTRYPGAYYIFQIKSGNEGEFYMROGPISTATVMTPRPKGPREIQIDLEMI 422
 DB 1144 GRAPAFAGDTISLTITKGNEGFYVTRRLNAYIGVSLQKSVLEPRDPAIDVEM 1197

RESULT 6
 A:Accession: A55184
 A:Title: fibulin-2 precursor - human
 N:Alternate names: protein DKFZp586A151.1
 C:Species: Homo sapiens (man)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Aug-2002
 C:Accession: A55184; T08744
 R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
 Genomics 22, 425-430, 1994
 A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the
 A:Reference number: A55184; MUID:95104855; PMID:7806230
 A:Accession: A55184
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1184 <ZHA>
 A:Cross-references: GB:X82494; NID:9575232; PIDN:CA57876.1; PID:9575233
 R:Wandut, R.; Heudner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z16471
 A:Accession: T08744
 A:Molecule type: mRNA
 A:Residues: 656-719, 'ODECLMGADCSRQPCVNTLGSFYCVNHTVLCADGTYILNAHKCYD', 720-853, 'T', 855
 A:Cross-references: EMBL:AD050095
 A:Experimental source: adult uterus; clone DKFZp586A1519
 C:Genetics:
 A:Gene: GDB:FBLN2
 A:Cross-references: GDB:293037; OMIM:135821
 A:Map position: 3p25-3p24
 A:Note: DKFZp586A151.1
 C:Superfamily: fibulin-2; EGF homology
 C:Keywords: alternative splicing; extracellular matrix
 F:1-27/Domain: signal sequence #status predicted <Sig>
 F:28-1184/Product: fibulin-2 protein #status predicted <Mat>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:36:12 ; Search time 18.0023 Seconds

(without alignments)
2392.374 Million cell updates/sec

Title: US-09-674-379A-13

Perfect score: 2533
Sequence: 1 MGIKRLITVITIALCLPSP.....INFRGSVRLRLRYVSQYF 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1109.5	43.8	493	2 JC5621	epidermal growth f
2	963.5	38.0	387	2 138449	extracellular prot
3	746	29.5	685	2 S78040	fibulin, splice fo
4	737	29.1	683	2 C36346	fibulin, splice fo
5	733.5	29.0	1221	2 A49457	fibulin-2 precursor
6	709.5	28.0	1184	2 A55184	fibulin-2 precursor
7	703.5	27.8	705	2 S34968	fibulin, splice fo
8	619	24.4	601	2 B36346	fibulin 1 precursor
9	606.5	23.9	689	2 T42760	fibulin, splice fo
10	599.5	23.7	712	2 T42990	fibulin 1, splice
11	596	23.5	589	2 T43210	fibulin-1D precurs
12	577.5	22.8	798	2 T22793	hypothetical prote
13	544	21.5	1394	2 A35626	transforming grow
14	532.5	21.0	3002	2 A47221	fibritillin 1 precu
15	531.5	21.0	2871	2 A55567	fibritillin 1 - bov
16	523	20.6	1712	2 A58261	masking protein pr
17	518.5	20.5	2871	2 A55624	fibritillin-1 precu
18	518.5	20.5	2907	2 A57278	fibritillin-2 precu
19	517.5	20.4	2918	2 A54105	latent transformin
20	480.5	19.0	1820	2 A55494	latent transformin
21	475	18.8	741	2 T46488	hypothetical prote
22	461	18.2	1251	2 A57293	latent transformin
23	459	18.1	1620	2 T27283	hypothetical prote
24	444	17.5	1574	2 T13954	MEGF6 protein - ra
25	392	15.5	886	2 A57172	probable hormone r
26	388	15.3	3507	2 T34513	hypothetical prote
27	372.5	14.7	1106	2 T18739	hypothetical prote
28	371.5	14.7	2471	2 A49128	cell-fate determin
29	365.5	14.4	1203	2 A49175	Notch B protein -

30	364	14.4	810	2 T10756	Nel-homolog protei
31	361	14.3	1081	2 T31329	receptor tyrosine
32	359.5	14.2	1964	2 T09059	notch4 - mouse
33	359	14.2	511	2 T17298	hypothetical prote
34	358	14.1	2555	2 A40043	notch protein homo
35	356	14.1	2703	1 A24420	notch protein - fr
36	355.5	14.0	2437	2 A42612	transmembrane prot
37	346	13.7	1064	2 A35844	fibropellin 1a - s
38	346	13.7	2524	2 A40136	notch protein - Af
39	344	13.6	2321	2 S78549	notch3 protein - h
40	340.5	13.4	673	2 A48089	growth arrest-spec
41	337.5	13.3	2318	2 S45306	notch 3 protein
42	337.5	13.3	2531	2 S18188	notch protein homo
43	337	13.2	2531	2 T31070	notch homolog - se
44	335.5	13.2	674	2 I55476	growth potentialin
45	335	13.2	2531	2 A46019	Notch-1 protein -

ALIGNMENTS

RESULT 1

JC5621 epidermal growth factor-like protein, T16 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 05-Nov-1999

C:Accession: JC5621

R:Ozaki, T.; Kondo, K.; Nakamura, Y.; Ichimiya, S.; Nakagawara, A.; Sakiyama, S.

Biochem. Biophys. Res. Commun. 237, 245-250, 1997

A:Title: Interaction of DA41, a DAN-binding protein, with the epidermal growth factor

A:Reference number: JC5621; MUID:97415782; PMID:9268694

A:Accession: JC5621

A:Molecule type: mRNA

A:Residues: 1-493 <OZA>

A:Cross-References: DDBJ:D89730; NID:q2429082; PIDN:BA42265.1; PID:q1023127; PID:q2

C:Comment: This protein plays a role in the regulation of cell growth by interacting

C:Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <Sig>

F:28-70,158-199,200-237,238-277,278-318,319-359/Region: epidermal growth factor-like

F:249/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 43.8%; Score 1109.5; DB 2; Length 493;

Best local similarity 42.3%; Pred. No. 1.1e-65;

Matches 206; Conservative 73; Mismatches 161; Indels 47; Gaps 5;

QY	7	ILTVTIALCLPSPENAGACCTNGRDLDROSQCDIDECRTPEACRGDMQVNGGY	66
DB	9	MTLALVKSQVTEETITTYOCTDGEWPDVROCKDIDCDIAPDCKGCMKCVNHGYG	68
QY	67	LCIPRTNVPYRGPSNPRSTPYS-----GPPAAPPISAP	102
DB	69	LCIPRTAOTIYNNEEDPOQETPAEASSGAATGTIAARSMATSGVTGGGFIATAVAGP	128
QY	103	NYPT-----ISRLICFRGYQMDSENOCVDVDECATDSHCNPTQ	142
DB	129	EVQGRNNFVIRBNPADPQRIIPSNPSHRIQCAAGYQSHNVCODIDECTSGHNRLDQ	188
QY	143	ICINTEGTYTSCIDGYMLBEGCQDIDECRI-GTCQQLCANVPGSYSCNGFTLNED	201
DB	189	VCINLRGFTCHCLPGLYGRGECVDIDECSSVPYCHQCQVTPGSGFYCQCNFGFLAAN	248
QY	202	GRSCQDVNECATENPCVQTVNTVGSFTCRCPGYELBEGVHCSMDRESEFSCQHE	261
DB	249	NYTCVDVINECASNOCACQCVITLISFTCCQNGVELSDRLNCEIDECRTSSYLQCYQ	308
QY	262	CVNPGYFSCSPRYILLDNRSQDINECHRNTCMLOQTCVMIQGGFKCIDIRCE	321
DB	309	CVNPGKFSKCMCPQGYQVY-RSKTCODINECTTNE-CDEDECMAYHGGFRKYPPNQ	366
QY	322	EPVLRISDNRCMAENPDCRQDPFTIYRDMVVGSRVADIEQMAATRYGAYVF	361
DB	367	DPVYLTSENRCVCPVSNMTCRDVPOSIVYKYVMNIRSRSVPDIFQIATTYANTINF	426

XX (ONCY) ONO PHARM CO LTD.
PA
XX
XX

PI Honjo T, Tashiro K, Nakamura T;
XX

DR WPI: 2000-038646/03.
XX

XX N-PSDB: AAZ39386, AAZ39387.
XX

PT Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma
XX

PS Claim 1; Page 54-56; 70pp: Japanese.
XX

CC The invention provides mouse polypeptides for treatment of diseases due
CC to abnormal proliferation of smooth muscle. The polypeptides can be
CC produced by standard recombinant methodology. The polypeptides can be
CC used according to their inhibition of the proliferation of vascular
CC smooth muscle cells, particularly in treating arteriosclerosis or re-
CC narrowing by vascular endothelial thickening after percutaneous
CC transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-
CC regulatory activity, cytokine activity, tissue generation/repairation
CC activity, actin/inhibitin activity, taxis and chemotaxis activity, blood
CC coagulation/thrombotic activity, receptor/ligand activity, cadonelin/
CC tumour metastasis inhibiting activity; tumor inhibition, and as nutrient.
CC The present sequence represents the protein of the invention which can be
CC used for modulating smooth muscle cell proliferation.
XX

S0 Sequence 461 AA:

Query Match 95.6%; Score 2302; DB 21; Length 461;
Best Local Similarity 94.8%; Pred. No. 2.5e-153;

Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

QY 1 OCTNFGDLDROSGGQCCLDIDECRRTPACRGDMCMVONGGYLCIPRTNPVYRGPSNPS 60
DB |||||||
39 OCTNFGDLDROSGGQCCLDIDECRRTPACRGDMCMVONGGYLCIPRTNPVYRGPSNPS 98
QY 61 TPYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDECATDSHCNPTQICI 120
DB |||||||
99 TSYSGPYPAAPPLSAPNTPTISRLICRGYOMDESNOCVDECATDSHCNPTQICI 158
QY 121 NTEGGYTCSTGCVWLLLEGQCCLDIDECRRGYCOQLCANVPGSYSCTCNPFTLNEDGRSC 180
DB |||||||
159 NTEGGYTCSTGCVWLLLEGQCCLDIDECRRGYCOQLCANVPGSYSCTCNPFTLNEDGRSC 218
QY 181 ODVNECATPNPCVQVNTYGSFICRCDDRGYELFEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB |||||||
219 ODVNECATPNPCVQVNTYGSFICRCDDRGYELFEDGVHCSMDDECSFSEFLCOHECVNQ 278
QY 241 PGTYFCSCPFGYLLDLDNRSCODINCEHNRHTCNLQOTCYNLOGGFKCIDPICEEPLYL 300
DB |||||||
279 PGTYFCSCPFGYLLDLDNRSCODINCEHNRHTCTSLQTCYNLOGGFKCIDPICEEPLYL 338
QY 301 RISDNRCMCPAENPCGRDQPTFLYRDMDVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB |||||||
339 LIGENRRCMCPAENPCGRDQPTFLYRDMDVSGRSVPADIFOMQATTRYPGAYYIFQIKS 398
QY 361 GNEGREFYMRQGTSPATLWTRRPIRGPREIOQDLEMTVNTVYINFGSSYIRLRIYYSQ 420
DB |||||||
399 GNEGREFYMRQGTSPATLWTRRPIRGPREIOQDLEMTVNTVYINFGSSYIRLRIYYSQ 458
QY 421 YPF 423
DB |||
459 YPF 461

Search completed: July 3, 2003, 17:54:48
Job time : 30.5959 secs

CC can be coeluted with agents that bind to these polypeptides, resulting in the death of the cells. The polynucleotides encoding these polypeptides are useful in the recombinant production of the polypeptides, as a hybridisation probe to screen libraries to isolate homologous sequences, or to map the gene. They may also be used for analysing genetic disorders, and to produce transgenic animals which are useful for the development and screening of therapeutically useful reagents. The polynucleotides can also be used in gene therapy e.g. to replace a defective gene.

SQ Sequence 448 AA;

Query Match	99.88;	Score 2401;	DB 22;	Length 448;
Best Local Similarity	99.88;	Pred. No. 2.8e-160;		
Matches 422; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	QCTNGEDLBROSQCGLDIDECRTTPEACRGDMKCVNONGYLC1PRTNPVYRGYSNPS	60
Db	26	QCTNGEDLBROSQCGLDIDECRTTPEACRGDMKCVNONGYLC1PRTNPVYRGYSNPS	85
QY	61	TFYSGYPAAAPPLSAPNPYTIISRPLLCRGFYQADESNOCVDYDECATDSHOCNTOJCI	12
Db	86	TFYSGYPAAAPPLSAPNPYTIISRPLLCRGFYQADESNOCVDYDECATDSHOCNTOJCI	14
QY	121	NTEGGTYTCTGTYMLLEGQCLDIDECRYCYCOOLCANPVSYSCTCNPGFTLMDGRSC	18
Db	146	NTEGGTYTCTGTYMLLEGQCLDIDECRYCYCOOLCANPVSYSCTCNPGFTLMDGRSC	20
QY	181	QDVNFCATNPCCVQTCVNTYGSFTRCQDDPGYTELEEDGVHSDMECSFSEFLCOHECVNQ	24
Db	206	QDVNFCATNPCCVQTCVNTYGSFTRCQDDPGYTELEEDGVHSDMECSFSEFLCOHECVNQ	26
QY	241	PGTYCSCPGYIILLDNRSCODINECEHNNHCNLLQOTCVNLQGFKECIPICEEPEYL	30
Db	266	PGTYCSCPGYIILLDNRSCODINECEHNNHCNLLQOTCVNLQGFKECIPICEEPEYL	32
QY	301	RISDRMCAPANPCCRODPFTILYRDMOVVSRSVPADIFQOATTRYPGAYYIFQIKS	36
Db	326	RISDRMCAPANPCCRODPFTILYRDMOVVSRSVPADIFQOATTRYPGAYYIFQIKS	38
QY	361	GNEGHEFYMRQTPISATLYMTRPIKGPKEIOLDLEMTYNTYINFGSSVYIRLRIYVSQ	42
Db	386	GNEGHEFYMRQTPISATLYMTRPIKGPKEIOLDLEMTYNTYINFGSSVYIRLRIYVSQ	44
QY	421	YPF 423	
Db	446	YPF 448	

RESULT 11
AAAY56751

AAV56751 standard; Protein; 423 AA.

AC AAY56751;

22-FEB-2000 (first entry)

XX muscle proliferation modulating protein mature sequence.

KW cell proliferation; vasculature; smooth muscle cell; arteriosclerosis; pTGA,
KW endothelial thickening; percutaneous transluminal coronary angioplasty;
KW myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis;
KW actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse;
KW metastasis; nutrient.

OS Mus musculus.

PN W09955863-A1.

PD 04-NOV-1999

28-APR-1999; 99WO-JP02283.

PR 28-APR-1998; 98JP-0119731.
YY

PA (ONOX) ONO PHARM CO LTD.
YY

PI Honjo T, Tashiro K, Nakamura T;
 YX

DR WPI; 2000-038646/03.
DB N-PEDD; 2000-038646/03.

DR N-PSDB; AAZ39385.
YY

Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

PS Claim 1; Page 51-53; 70pp; Japanese.
 YX

The invention provides mouse polypeptides for treatment of diseases due to abnormal proliferation of smooth muscle. The polypeptides can be produced by standard recombinant methodology. The polypeptides can be used according to their inhibitory of the proliferation of vascular smooth muscle cells, particularly in treating atherosclerosis or re-narrowing by vascular endothelial thickening after percutaneous transluminal coronary angioplasty (PTCA), or myoma, hematopoietic cell-regulatory activity, cytokine activity, tissue generation/repair activity, actin/inhibitin activity, taxis and chemotaxis activity, blood coagulation/thrombotic activity, receptor/ligand activity, cadherin/ tumour metastasis inhibiting activity, tumor inhibition, and as nutrient. The present sequence represents the mature protein of the invention which can be used for modulating smooth muscle cell proliferation.

Sequence 423 AA;

Query Match	95.6%	Score 2302;	DB 21;	Length 423;
Best Local Similarity	94.8%	Pred. No. 2.2e-153;		
Matches 401; Conservative	9;	Mismatches 13;	Indels 0;	Gaps 0

QY	1	OCNNGSDLDROSQCCLDIDECRTIPEACGDMCYNONGYLCIPIRTNPVYRGPSNYS	60
Db	1	OCNNGSDLDROSQCCLDIDECRTIPEACGDMCYNONGYLCIPIRTNPVYRGPSNYS	60
QY	61	TPYSGPYPAAPAEPLSAPNPPTISRPPLICFEGYQOMESNOCYDVBECADSHOQNPOTICI	120
Db	61	TPYSGPYPAAPAEPLSAPNPPTISRPPLICFEGYQOMESNOCYDVBECADSHOQNPOTICI	120
QY	121	MTEGGTGCCTGTGYMLLBESQCLDIDECRYGVCQQLCANVPASYSCTCNPGFTLNEDGASC	180
Db	121	MTEGGTGCCTGTGYMLLBESQCLDIDECRYGVCQQLCANVPASYSCTCNPGFTLNEDGASC	180
QY	181	ODVNECATENPCVOTCVMTYSGSFICRCPGVELEDEGTHGSDMECSSEFLCQHECVNQ	240
Db	181	QDVNECEIENPCVQTCVMTYSGSFICRCPGVELEDEGTHGSDMECSSEFLCQHECVNQ	240
QY	241	PCTYTCSCPBGYIILLDNRSQCQDINECEHRNHTCNLAOTCYNLGGFEKCIDPIRICEEPLY	300
Db	241	PESTYTCSCPBGYIILLDNRSQCQDINECEHRNHTCSLTCTCNLGGFEKCIDPIRICEEPLY	300
QY	301	RTSDNRGCMCAENPCRDQPFITILYRDMDDVYSGSVPADIFQOMATTRYPGAYITFOIKS	360
Db	301	LIGENRKCPCAPBHHSCRDQPFITILYRDMDDVYSGSVPADIFQOMATTRYPGAYITFOIKS	360
QY	361	GENEGREFMYRQTPGISATLVMTRPILKGRFQIOLDLEMTVNTVYNFNGSSVYIRLRIYVQ	420
Db	361	GENEGREFMYRQTPGISATLVMTRPILKGRFQIOLDLEMTVNTVYNFNGSSVYIRLRIYVQ	420
QY	421	YPF 423	
Db	421	YPF 423	
QY	421	YPF 423	
Db	421	YPF 423	

RESULT 12

AAAY56753

AA56753 standard; Protein; 423 AA.

AC AY56753;

XX

CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adenoma, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 448 AA:

Query Match 99.88; Score 2401; DB 22; Length 448;
Best Local Similarity 99.88; Pred. No. 2,8e+160;
Matches 442; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMKCVNNGGYLCIPRTNPPYRGPSNPPS 60
26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMKCVNNGGYLCIPRTNPPYRGPSNPPS 85
61 TPYSGPYPAAPLAPLAPNPTISRPLICRGYQMDENQCVADDECATDSHQCNPFOICI 120
86 TPYSGPYPAAPLAPLAPNPTISRPLICRGYQMDENQCVADDECATDSHQCNPFOICI 145
121 NTEGGYTCGCTDGVNMLEGQCLDIDECRYGCOQLCANVPSSYSTCNPFTLNEDGRSC 180
146 NTEGGYTCGCTDGVNMLEGQCLDIDECRYGCOQLCANVPSSYSTCNPFTLNEDGRSC 205
181 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGHCSDMDECSFEFLQHECVNQ 240
206 QDVNECATENPCVQTCVNTYGSFICRCDPGYELEDGHCSDMDECSFEFLQHECVNQ 265
241 PGTYFSCPPGYLLDNRSCQDINECEHNNHCNLOQTCYNLOGGFKCIPDICEPEPYL 300
266 PGTYFSCPPGYLLDNRSCQDINECEHNNHCNLOQTCYNLOGGFKCIPDICEPEPYL 325
301 RISDNMCMCPAEMPGCRDOPFTLLYRDMVYSGRSVPADIDOMATTYPPAYYFQIKS 360
326 RISDNMCMCPAEMPGCRDOPFTLLYRDMVYSGRSVPADIDOMATTYPPAYYFQIKS 385
361 GNEGREFYKQTPISATLVMTPIKGPRIQDLEMTVTNVTINFRGSSVIRLRIYYSQ 420
386 GNEGREFYKQTPISATLVMTPIKGPRIQDLEMTVTNVTINFRGSSVIRLRIYYSQ 445
421 YPF 423
446 YPF 448

RESULT 10

ID AAB31183 standard; Protein; 448 AA.

XX AAB31183;

20-APR-2001 (first entry)

XX Amino acid sequence of human polypeptide PRO210.

XX Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
XX PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1308;
XX PRO1183; PRO1272; PRO1419; PRO4999; PRO1170; PRO248; PRO353; PRO1318;
XX PRO1600; PRO9940; PRO333; PRO301; PRO187; PRO337; PRO411; PRO4356;
XX PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO350; PRO630;
XX PRO6309; cell death; genetic disorder; transgenic animal; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..25 /note- "signal peptide"

FT Modified-site 21..27 /note- "N-myristoylation site"

FT Binding-site 54..57 /note- "cell attachment site"

FT Modified-site 64..70

FT /note- "N-myristoylation site"
FT 144..156 /note- "aspartic acid and asparagine hydroxylation site"
FT Modified-site 149..155 /note- "N-myristoylation site"
FT Modified-site 186..192 /note- "N-myristoylation site"
FT Modified-site 226..232 /note- "N-myristoylation site"
FT Modified-site 242..248 /note- "N-myristoylation site"
FT Modified-site 267..273 /note- "N-myristoylation site"
FT Modified-site 283..287 /note- "N-glycosylation site"
FT Modified-site 296..300 /note- "N-glycosylation site"
FT Modified-site 310..316 /note- "N-myristoylation site"

W0200077037-A2.

XX 21-DEC-2000.
XX 22-MAY-2000; 2000MO-US14042.
XX 15-JUN-1999; 99US-0139695.
XX 20-JUL-1999; 99US-0145070.
XX 26-JUL-1999; 99US-0145698.
XX 17-AUG-1999; 99US-0149396.
XX 01-SEP-1999; 99MO-US20111.
XX 08-SEP-1999; 99MO-US20594.
XX 15-SEP-1999; 99MO-US21090.
XX 15-SEP-1999; 99MO-US21547.
XX 30-NOV-1999; 99MO-US28313.
XX 01-DEC-1999; 99MO-US28301.
XX 02-DEC-1999; 99MO-US28565.
XX 07-DEC-1999; 99US-0169495.
XX 05-JAN-2000; 2000MO-US00219.
XX 18-FEB-2000; 2000MO-US04341.
XX 18-FEB-2000; 2000MO-US04342.
XX 22-FEB-2000; 2000MO-US04414.
XX 01-MAR-2000; 2000MO-US05601.
XX 02-MAR-2000; 2000MO-US05841.
XX 20-MAR-2000; 2000MO-US07377.
XX 30-MAR-2000; 2000MO-US08439.
XX 15-MAY-2000; 2000MO-US13358.
XX 17-MAY-2000; 2000MO-US13705.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
XX Ferrara N, Fong S, Gao W, Geider H, Gerritsen ME, Goddard A;
XX Godowski PJ, Gurney AL, Kljavin JV, Mather JP, Napier MA, Pan J;
XX Paoni NF, Roy MA, Stewart JA, Tumas D, Watanabe CK, Williams PW;
XX Wood WI, Zhang Z;

DR WPI: 2001-050091/06.

DR N-PSDB; AAC69668.

XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a
XX transmembrane polypeptide is useful for gene therapy and identification
XX of related polypeptides -

PS Claim 12; Fig 10; 24pp; English.

XX The present sequence represents a human secreted and transmembrane
XX polypeptide. The specification describes human polypeptides, designated
XX PRO196, PRO444, PRO183, PRO210, PRO215, PRO217, PRO242, PRO288,
XX PRO365, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
XX PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
XX PRO337, PRO411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
XX PRO6004, PRO350, PRO265 and PRO6309. The biological activity of cells

CC Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic
 CC thyroiditis), diabetes mellitus, immune-mediated renal disease
 CC (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,
 CC idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic
 CC inflammatory demyelinating polyneuropathy, infectious hepatitis
 CC (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune
 CC chronic active hepatitis, primary biliary cirrhosis, granulomatous
 CC hepatitis, and sclerosing cholangitis, inflammatory bowel disease
 CC (ulcerative colitis: Crohn's disease), gluten-sensitive enteropathy, and
 CC Whipple's disease. Autoimmune or immune-mediated skin diseases including
 CC bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,
 CC asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,
 CC urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,
 CC hypersensitivity pneumonitis, and transplantation associated diseases
 CC (graft rejection, and graft-versus-host-disease). (i) Its (ant)agonists
 CC or fragment can also be used as an adjuvant in treatment of tumors.
 CC Antibodies against (i) can also be used for diagnosing such diseases.
 CC This sequence represents a human EGF-like homologue encoded by cDNA clone
 CC DN32279 which is described in the invention.

XX Sequence 448 AA;

Query Match 99.8%; Score 2401; DB 20; Length 448;

Best Local Similarity 99.8%; Pred. No. 2.8e-160;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QCTNGFDLRQSGCCLDIDECRTPEACRCGDMCVNNGGYLCIPRTNRYRGPSNYS 60
 DB 26 OCTNGFDLRQSGCCLDIDECRTPEACRCGDMCVNNGGYLCIPRTNRYRGPSNYS 85
 QY 61 TPYSGPAPAPPLSAPNYPITISRPILCRGYQMDSESNQCVNDECATDSHCNPTQICI 120
 DB 86 TPYSGPAPAPPLSAPNYPITISRPILCRGYQMDSESNQCVNDECATDSHCNPTQICI 145
 QY 121 MTEGGYTCTDGYWMLLEGSCCLDIDECRYGCOQLCANVGSYCTCPNPGFTLNEDESRSC 180
 DB 146 MTEGGYTCTDGYWMLLEGSCCLDIDECRYGCOQLCANVGSYCTCPNPGFTLNEDESRSC 205
 QY 181 ODVNECATENPCVQTCVNTYGSFTCRCPGPELLEEDGVHCSMDDECSFSEFLCHQECVNO 240
 DB 206 ODVNECATENPCVQTCVNTYGSFTCRCPGPELLEEDGVHCSMDDECSFSEFLCHQECVNO 265
 QY 241 PGTYFPCSPPGYLLDDNRSQDINCEHNRHNCNLOQTCYNLOGGKCKDPTICEPPL 300
 DB 266 PGTYFPCSPPGYLLDDNRSQDINCEHNRHNCNLOQTCYNLOGGKCKDPTICEPPL 325
 Y 301 RLSDRNCKPAENPGCRDOPFTLLYRDMOVSGRSVPADIFQMATRYPGAYYIFQIKS 360
 DB 326 RLSDRNCKPAENPGCRDOPFTLLYRDMOVSGRSVPADIFQMATRYPGAYYIFQIKS 385
 QY 361 GNEGREFFYRGTPTISATLVTMRPTKGRREIOLDEMTIVNTYINRGSSVILRLIYVSQ 420
 DB 386 GNEGREFFYRGTPTISATLVTMRPTKGRREIOLDEMTIVNTYINRGSSVILRLIYVSQ 445
 QY 421 YPF 423
 DB 446 YPF 448

RESULT 9

AAU29227 standard: Protein; 448 AA.

AAU29227:

18-DEC-2001 (first entry)

Human PRO polypeptide sequence #204.

XX PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX OS Homo sapiens.
 XX PN WO200168848-A2.
 XX PD 20-SEP-2001.
 XX PF 28-FEB-2001; 2001WO-US06520.
 XX PR 01-MAR-2000; 2000WO-US05601.
 XX PR 02-MAR-2000; 2000WO-US05841.
 XX PR 03-MAR-2000; 2000US-187202P.
 XX PR 04-MAR-2000; 2000US-186968P.
 XX PR 14-MAR-2000; 2000US-189320P.
 XX PR 15-MAR-2000; 2000US-189328P.
 XX PR 21-MAR-2000; 2000WO-US06884.
 XX PR 21-MAR-2000; 2000US-190828P.
 XX PR 21-MAR-2000; 2000US-191007P.
 XX PR 21-MAR-2000; 2000US-191048P.
 XX PR 21-MAR-2000; 2000US-191314P.
 XX PR 28-MAR-2000; 2000US-192655P.
 XX PR 29-MAR-2000; 2000US-193032P.
 XX PR 29-MAR-2000; 2000US-193053P.
 XX PR 30-MAR-2000; 2000WO-US08439.
 XX PR 04-APR-2000; 2000US-194449P.
 XX PR 04-APR-2000; 2000US-194647P.
 XX PR 11-APR-2000; 2000US-195975P.
 XX PR 11-APR-2000; 2000US-196000P.
 XX PR 11-APR-2000; 2000US-196187P.
 XX PR 11-APR-2000; 2000US-196690P.
 XX PR 18-APR-2000; 2000US-196820P.
 XX PR 18-APR-2000; 2000US-198121P.
 XX PR 18-APR-2000; 2000US-198585P.
 XX PR 25-APR-2000; 2000US-199397P.
 XX PR 25-APR-2000; 2000US-199550P.
 XX PR 25-APR-2000; 2000US-199654P.
 XX PR 03-MAY-2000; 2000US-201516P.
 XX PR 17-MAY-2000; 2000WO-US13705.
 XX PR 22-MAY-2000; 2000WO-US14042.
 XX PR 30-MAY-2000; 2000WO-US14941.
 XX PR 02-JUN-2000; 2000WO-US15264.
 XX PR 05-JUN-2000; 2000US-209832P.
 XX PR 28-JUL-2000; 2000WO-US20710.
 XX PR 22-AUG-2000; 2000US-064484P.
 XX PR 24-AUG-2000; 2000WO-US33328.
 XX PR 08-NOV-2000; 2000WO-US30952.
 XX PR 01-DEC-2000; 2000WO-US32678.
 XX PR 20-DEC-2000; 2000WO-US34956.
 XX PA (GENTH) GENENTECH INC.
 XX PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 XX PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX DR WPI; 2001-602746/68.
 XX DR N-PSDB; AAS46128.
 XX PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 XX PT presence of tumours, such as prostate and breast tumours, in mammals and
 XX PT to screen for modulators of the compounds -
 XX PS Claim 11: Fig 408; 774pp; English.
 XX PS Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
 XX CC The PRO polypeptides and their associated nucleic acids can be used to
 XX CC detect the presence of a tumour in a mammal by comparing the level of
 XX CC expression of a PRO polypeptide in a test sample of cells from the animal
 XX CC and a control sample of normal cells, whereby a higher level of
 XX CC expression in the test sample indicates the presence of a tumour in the
 XX CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 XX CC and rabbits but are preferably human. The polypeptides can be used to
 XX CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
 XX CC when contacted with it. A specific polypeptide can be used to stimulate

Db 266 PGTYFCSCPFGYILLDNRSQCQDINECEHNRHTCNLQOTCVNLGGFKCIDPRTNCEPYL 325
 QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 360
 Db 326 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 385
 QY 361 GNEGREFFMRQTPISATLVMTTRPIKGRREIQLDLEMTVNTVINRGSSVIRLRIVSQ 420
 Db 386 GNEGREFFMRQTPISATLVMTTRPIKGRREIQLDLEMTVNTVINRGSSVIRLRIVSQ 445
 QY 421 YPF 423
 Db 446 YPF 448

RESULT 6

AA093573 standard; Protein; 448 AA.

AC AAM93573:
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide, SEQ ID NO: 3357.
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 OS Homo sapiens.
 PN EP1130094-A2.
 PD 05-SEP-2001.
 PF 07-JUL-2000; 2000EP-0114089.
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 PA (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-524255/58.
 DR N-PSDB: AAK94505.
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -

PS Claim 8: SEQ ID NO 3357; 1380bp + sequence listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 SO Sequence 448 AA:

Query Match 100.0%; Score 2407; DB 22; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.le-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QCTNGFDLDROSGCCLDIDECRTIPACRGDMCVNONGYLICIPRTNPPYRGPSNPYS 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 26 QCTNGFDLDROSGCCLDIDECRTIPACRGDMCVNONGYLICIPRTNPPYRGPSNPYS 85
 QY 61 TPYSGPYPAAPPLSAPNYPPLISRPILCRREFYOMDESQCVQVDEBCADSHQCNTOICI 120
 Db 86 TPYSGPYPAAPPLSAPNYPPLISRPILCRREFYOMDESQCVQVDEBCADSHQCNTOICI 145
 QY 121 NTEGGYTCSTGQWMLLEGGCCLDIDECRYGCOQLCANVPSSYCTCNPFTLNEDGRSC 180
 Db 146 NTEGGYTCSTGQWMLLEGGCCLDIDECRYGCOQLCANVPSSYCTCNPFTLNEDGRSC 205
 QY 181 QDVNCEATENPCVQTCVNTGSGFICRCDDPGYLEEDGVHCSMDCECSFEFLCOHECVNQ 240
 Db 206 QDVNCEATENPCVQTCVNTGSGFICRCDDPGYLEEDGVHCSMDCECSFEFLCOHECVNQ 265
 QY 241 PGTYFCSCPFGYILLDNRSQCQDINECEHNRHTCNLQOTCVNLGGFKCIDPRTNCEPYL 300
 Db 266 PGTYFCSCPFGYILLDNRSQCQDINECEHNRHTCNLQOTCVNLGGFKCIDPRTNCEPYL 325
 QY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 360
 Db 326 RISDNRCMCPAENPCGRDOPFTILYRDMDVSGRSVPADIFQMOATRRYPGAYYIFQIKS 385
 QY 361 GNEGREFFMRQTPISATLVMTTRPIKGRREIQLDLEMTVNTVINRGSSVIRLRIVSQ 420
 Db 386 GNEGREFFMRQTPISATLVMTTRPIKGRREIQLDLEMTVNTVINRGSSVIRLRIVSQ 445
 QY 421 YPF 423
 Db 446 YPF 448

RESULT 7

AA075494 standard; Protein; 448 AA.

AC AA075494:
 DT 23-APR-2002 (first entry)
 DE Human extracellular protein-1like/EGF-like protein, EECF.
 KW Human; extracellular protein-1like protein; EGF-1like;
 KW protein; epidermal growth factor; EGF; ATCC 97285; gene therapy;
 KW vascular smooth muscle cell proliferation; Marfan syndrome;
 KW wound healing; neurological trauma; acquired immunodeficiency syndrome;
 KW AIDS-related dementia; ocular disorder; kidney disorder; liver disorder;
 KW hair follicle growth promotion; burn; ulcer; corneal incision;
 KW corneal inflammation; neoplasm; psoriasis.
 OS Homo sapiens.
 PH Key

FT Location/Qualifiers
 FT 1..25 /label= Signal_peptide
 FT 26..48 /label= Mature_EEGF
 FT /note= "This region is specifically claimed in claim 10"
 FT 112..153 /label= EGF_1_domain
 FT /note= "This domain is specifically claimed in claim 10"
 FT 154..190 /label= EGF_2_domain
 FT /note= "This domain is specifically claimed in claim 10"
 FT 191..230 /label= EGF_3_domain
 FT /note= "This domain is specifically claimed in claim 10"
 FT 231..271 /label= EGF_4_domain
 FT /note= "This domain is specifically claimed in claim 10"
 FT 272..314 /label= EGF_5_domain
 FT /note= "This domain is specifically claimed in claim 10"
 FT Domain
 FT Domain
 FT Domain
 FT Domain

PT smooth muscle cell growth, vasculogenesis, restenosis or
 XX atherosclerosis
 PS
 XX Claim 4; Page 46-47; 49pp; English.

CC This is the amino acid sequence of the human secreted protein AK647. The
 CC polynucleotide sequence was obtained from a human foetal kidney cDNA
 CC library. AK647 homologues in chicks and rodents are involved in aortic
 CC tissue development. The spatial and temporal distribution of AK647
 CC indicated that it acts as an a modulator of smooth muscle cells in
 CC vasculogenesis during embryonic development. The primary structure of
 CC AK647 consists of multiple EGF domains. The AK647 protein can be used as
 CC a nutritional source or supplement. The protein shows both inhibitory and
 CC inducing, cytokine, cell proliferation and cell differentiation activity.
 CC The protein may also be used in the treatment of immune deficiencies and
 CC disorders, including severe combined immunodeficiency (SCID), HIV and
 CC other viral, bacterial and fungal infections. Regulation of immune
 CC responses may also be carried out by the AK647 protein. Other uses of the
 CC protein include a role in the regulation of haematopoiesis and in the
 CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
 CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
 CC as well as for wound healing and in the treatment of ulcers and burns.
 CC The polynucleotides and proteins can be used for preventing, treating or
 CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
 CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
 CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
 CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activity. AK647 specific
 CC antibodies can be used for promoting smooth muscle cell growth or
 CC vasculogenesis. The proteins and polynucleotides can also be used for
 CC detection, diagnosis and drug screening.

SQ Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.1e-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDGSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPPVYRGPSNPS 60
 DB 26 QCTNGFDLDGSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPPVYRGPSNPS 85
 QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 120
 DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 145
 QY 121 NTGSGYTCSTGDTGWLLEGGCCLDIDECRYGCGQQLCAVPSYSTCNPNGFTLNEDGRSC 180
 DB 146 NTGSGYTCSTGDTGWLLEGGCCLDIDECRYGCGQQLCAVPSYSTCNPNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEEDGVHCSMDSCSFSEFLCOHECVNQ 240
 DB 206 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEEDGVHCSMDSCSFSEFLCOHECVNQ 265
 QY 241 PGTYFSCPPGYITLLDNRSCQDINECEHRNHTCNLOOTCYNLGGFKCIDIRCEEPYL 300
 DB 266 PGTYFSCPPGYITLLDNRSCQDINECEHRNHTCNLOOTCYNLGGFKCIDIRCEEPYL 325
 QY 301 RISDNMCMCAENPGCDOPFTILYRDMVYSGRSVPADIFOMATTRYPGAYYIFQIKS 360
 DB 326 RISDNMCMCAENPGCDOPFTILYRDMVYSGRSVPADIFOMATTRYPGAYYIFQIKS 385
 QY 361 GNGRGREYMRQTPISATLVWTRPIKGRFETOLDLEMTVNTVINFRGSSVRLRLIYVSO 420
 DB 386 GNGRGREYMRQTPISATLVWTRPIKGRFETOLDLEMTVNTVINFRGSSVRLRLIYVSO 445
 QY 421 YPF 423
 DB 446 YPF 448

RESULT 5
 AAY54989

ID AAY54989 standard; Protein: 448 AA.

XX AC AAY54989;

XX DT 15-FEB-2000 (first entry)

XX DE Full length human A55 protein sequence.

XX KW A55 protein; human; smooth muscle proliferation; tissue generation;

XX KW vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;

XX KW percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;

XX KW actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;

XX OS Homo sapiens.

XX PN WO955864-A1.

XX PD 04-NOV-1999.

XX PF 28-APR-1999; 99WO-JP02284.

XX PR 28-APR-1998; 98JP-0119731.

XX PA (ONOX) ONO PHARM CO LTD.

XX PI Honjo T, Tashiro K, Nakamura T;

XX DR WPI; 2000-038647/03.

XX DR N-PSDB; AAZ40027.

XX PT Novel human polypeptides for treatment of, e.g. arteriosclerosis and

XX PS myoma

XX Claim 1; Page 76-80; 87pp; Japanese.

CC This sequence is the human A55 protein of the invention. The protein
 CC can be used for the treatment of diseases due to abnormal proliferation
 CC of smooth muscle. The polypeptides can be used according to their inhibition
 CC of the proliferation of vascular smooth muscle cells, particularly in
 CC treating arteriosclerosis or re-narrowing by vascular endothelial
 CC thickening after percutaneous transluminal coronary angioplasty (PTCA),
 CC or myoma, haematopoietic cell-regulatory activity, cytokine activity,
 CC tissue generation/repairation activity, actin/inhibin activity, taxis
 CC and chemotaxis activity, blood coagulation/thrombotic activity,
 CC receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
 CC tumour inhibition, and as a nutrient.

SQ Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 21; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.1e-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDGSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPPVYRGPSNPS 60
 DB 26 QCTNGFDLDGSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPPVYRGPSNPS 85
 QY 61 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 120
 DB 86 TPYSGPYPAAPPLSAPNPTISRPILCRFGYQMDENOCVDVDECATDSHOCNPTQICI 145
 QY 121 NTGSGYTCSTGDTGWLLEGGCCLDIDECRYGCGQQLCAVPSYSTCNPNGFTLNEDGRSC 180
 DB 146 NTGSGYTCSTGDTGWLLEGGCCLDIDECRYGCGQQLCAVPSYSTCNPNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEEDGVHCSMDSCSFSEFLCOHECVNQ 240
 DB 206 QDVNECATENPCVQTCVNTGSGFICRCDPGYLEEDGVHCSMDSCSFSEFLCOHECVNQ 265
 QY 241 PGTYFSCPPGYITLLDNRSCQDINECEHRNHTCNLOOTCYNLGGFKCIDIRCEEPYL 300
 DB 446 YPF 448

```

QY 61 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYQMDSESNOCVDVDECATDSHCNPOTICI 120
DB 86 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYQMDSESNOCVDVDECATDSHCNPOTICI 145
QY 121 NTEGGYTCTGCTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCTGCTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPGFTLNEDGRSC 205
QY 181 QDVNCAATENPCVOTCVNTYGSFICRCDPGYELEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNCAATENPCVOTCVNTYGSFICRCDPGYELEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCSPGTYILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 300
DB 266 PGTYFSCSPGTYILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 325
QY 301 RISDNRCMCPAENPGCGRDQPFITILYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCGRDQPFITILYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLVMTPIKGPREDIOLDEMTVNTVINFSGSVIRLRITYSQ 420
DB 386 GNEGREFYMRQTPISATLVMTPIKGPREDIOLDEMTVNTVINFSGSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

RESULT 3

AAW94281 ID AAW94281 standard; Protein: 448 AA.

07-MAY-1999 (first entry)

Human extracellular matrix protein (ECMP)-1.

Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;

Immune disorder; human.

Homo sapiens.

W09900410-A2.

07-JAN-1999.

23-JUN-1998; 98MO-US13012.

27-JUN-1997; 97US-0884072.

(INCY-) INCYTE PHARM INC.

Bandman O, Corley NC, Guegler KJ;

WPI: 1999-095674/08.

N-PSDB: AAX05359.

New polynucleotide encoding extracellular matrix protein, ECMP-1 -
useful in the diagnosis, prevention and treatment of immune
disorders and cancer

Claim 1; Fig 1A-G; 79pp; English.

This represents a human extracellular matrix protein (ECMP)-1. Host
cells containing a vector comprising the ECMP-1 nucleic acid are used
for the recombinant production of the protein. ECMP-1 and its
(ant)agonists, are useful in the diagnosis, prevention, and treatment
of cancer and immune disorders.

Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 20; Length 448;
Best Local Similarity 100.0%; Pred. No. 1,1e-160;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 OCTGFPDLDRSGGCLDIDECRTIPEACRGCMCVNONGYLICIPRTNPVYRGPSNPYS 60
DB 26 OCTGFPDLDRSGGCLDIDECRTIPEACRGCMCVNONGYLICIPRTNPVYRGPSNPYS 85
QY 61 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYQMDSESNOCVDVDECATDSHCNPOTICI 120
DB 86 TPYSGPYPAAPPLSAPNYPITISRPILCRFGYQMDSESNOCVDVDECATDSHCNPOTICI 145
QY 121 NTEGGYTCTGCTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCTGCTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPGFTLNEDGRSC 205
QY 181 QDVNCAATENPCVOTCVNTYGSFICRCDPGYELEDGVHCSMDDECSFSEFLCOHECVNQ 240
DB 206 QDVNCAATENPCVOTCVNTYGSFICRCDPGYELEDGVHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCSPGTYILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 300
DB 266 PGTYFSCSPGTYILLDDNRSQCODINECEHNNHTCNLOQTCYNLOGGFKCIDPICEPYL 325
QY 301 RISDNRCMCPAENPGCGRDQPFITILYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCGRDQPFITILYRDMDVYSGRSVPADIFQMOATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLVMTPIKGPREDIOLDEMTVNTVINFSGSVIRLRITYSQ 420
DB 386 GNEGREFYMRQTPISATLVMTPIKGPREDIOLDEMTVNTVINFSGSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

RESULT 4

AAV57058 ID AAV57058 standard; Protein: 448 AA.

AAV57058;

21-FEB-2000 (first entry)

Amino acid sequence of the human secreted protein AK647.

AK647; aortic tissue development; smooth muscle cell modulator; SCIP;
nutritional supplement; vasculogenesis; embryonic development; infection;
cytokine activity; cell proliferation; cell differentiation; detect; HIV;
immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
wound healing; restenosis; atherosclerosis; drug screen.

Homo sapiens.

W09960125-A2.

25-NOV-1999.

18-MAY-1999; 99WO-US10931.

19-MAY-1998; 98US-0081002.

21-MAY-1998; 98US-0083002.

(GENY) GENETICS INST INC.

Jacobs K, McCoy JM, Racie L, Lavallie E, Treacy M, Evans C;

Agostino M, Lu Z, Merberg D;

WPI: 2000-053298/04.

N-PSDB: AAZ39892.

Proteins, and their encoding polynucleotides, used for treating e.g.

FT Region /label= EGF-4
 FT 272..314
 FT /label= EGF-5
 XX
 XX MO9846746-A1.
 XX
 PD 22-OCT-1998.
 XX
 PF 11-APR-1997; 97MO-US06020.
 XX
 PK 11-APR-1997; 97MO-US06020.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI LI H, Olsen HS;
 XX
 DR WPI, 1998-568728/48.
 XX N-PSDB; AAV62432.
 XX
 PT New isolated extracellular/epidermal growth factor - used for
 PT regulating vascular smooth muscle cell proliferation, e.g. for
 PT enhancing neurological functions or treating neoplasia and other
 PT disorders.

PS Claim 10a; Fig 1A-D; 62pp; English.

XX This sequence represents a novel human extracellular/epidermal growth
 CC factor-like protein, EGF. This protein can be used to regulate
 CC vascular smooth muscle cell proliferation and for restoration or
 CC enhancement of neurological functions diminished as a result of other
 CC damaging pathologies such as AIDS dementia. The protein can also be used
 CC to treat senile dementia, ocular disorders such as corneal inflammation,
 CC for targeting tumour cells, for treating kidney disorders, for liver
 CC regeneration or treating liver dysfunction, for treating wounds including
 CC all cutaneous wounds, corneal wounds, and injuries to the
 CC epithelial-lined hollow organs of the body or resulting from trauma such
 CC as burns, abrasions and cuts as well as from surgical procedures such as
 CC surgical incisions and skin grafting. The polypeptides can also be used
 CC for treating chronic conditions, such as chronic ulcers, diabetic ulcers,
 CC other non-healing (trophic) conditions, to treat Marfan syndrome, to
 CC promote hair follicular development, to stimulate growth and
 CC differentiation of various epidermal and epithelial cells in vivo and in
 CC vitro and to stimulate embryogenesis. Antagonists to EGF can be used to
 CC treat neoplasia such as cancers or tumours, skin disorders such as
 CC psoriasis or corneal inflammation. The products can also be used for
 CC identifying EGF receptors, detection, diagnosis and drug screening.

XX Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 19; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.le-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 60
 DB OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 85
 QY 26 TYSGGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECAHDSHCNPTQICI 120
 DB 26 TYSGGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECAHDSHCNPTQICI 145
 QY 61 TYSGGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECAHDSHCNPTQICI 180
 DB 61 TYSGGYPAAPPLSAPNPTTISRPLCRGCGYQMDENOCVDNDECAHDSHCNPTQICI 205
 QY 121 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGYCOOLCANVPYSCTCNPFTLNEDGRSC 240
 DB 121 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGYCOOLCANVPYSCTCNPFTLNEDGRSC 265
 QY 146 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGYCOOLCANVPYSCTCNPFTLNEDGRSC 300
 DB 146 NTEGGYTCSTDGYWMLLEGQCLDIDECRYGYCOOLCANVPYSCTCNPFTLNEDGRSC 325
 QY 181 QDVNDECATENPCVQTCVNTYGSFICRCDPGYELLEDGSHCSMDDECSFSEFLCOHECVNO 240
 DB 181 QDVNDECATENPCVQTCVNTYGSFICRCDPGYELLEDGSHCSMDDECSFSEFLCOHECVNO 265
 QY 206 QDVNDECATENPCVQTCVNTYGSFICRCDPGYELLEDGSHCSMDDECSFSEFLCOHECVNO 300
 DB 206 QDVNDECATENPCVQTCVNTYGSFICRCDPGYELLEDGSHCSMDDECSFSEFLCOHECVNO 325
 QY 241 PGTYFSCSPRGYLLDLDNRSCODINEGHNHHCNLOOTCTNLOGGFRICDIPIRCEEYLL 300
 DB 241 PGTYFSCSPRGYLLDLDNRSCODINEGHNHHCNLOOTCTNLOGGFRICDIPIRCEEYLL 325

QY 301 RISDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSPADIFOMQATTRYPGAYYIFQJLS 360
 DB 326 RISDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSPADIFOMQATTRYPGAYYIFQJLS 385
 QY 361 GNEGREYMRQTPISATLVMTRPDKGPREIQLDLEMTVTNVTNFGSSVIRLRIVYSQ 420
 DB 386 GNEGREYMRQTPISATLVMTRPDKGPREIQLDLEMTVTNVTNFGSSVIRLRIVYSQ 445
 QY 421 YPF 423
 DB 446 YPF 448

RESULT 2

ID AAW95709 standard; Protein; 448 AA.

AAW95709;

21-JUN-1999 (first entry)

Homo sapiens fetal kidney clone AK647 secreted protein.

Secreted protein; fetal kidney.

Homo sapiens.

WO9900405-A1.

07-JAN-1999.

29-JUN-1998; 98MO-US13530.

30-JUN-1997; 97US-0885610.

(GEMV) GENETICS INST INC.

Agostino MJ, Evans C, Jacobs K, Lavallie ER, McCoy JM,

Merberg D, Racie LA, Treacy M;

WPI, 1999-095671/08.

N-PSDB; AAX07567.

New polynucleotides encoding secreted human proteins - are derived from foetal kidney or adult retina cDNA libraries, used as, e.g. potential vaccines

Claim 11; Pages 52-54; 76pp; English.

The sequence is that of a secreted protein from a human fetal kidney clone AK296. Such a sequence is predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting data is given. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. It is also stated to be useful for gene therapy.

Sequence 448 AA;

Query Match 100.0%; Score 2407; DB 20; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.le-160; Indels 0; Gaps 0;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 60
 DB OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVNGGGLCIPRTNPVYRGPSNYS 85

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 11:59:37 ; Search time 30.5959 Seconds
(without alignments)
1842.243 Million cell updates/sec

Title: US-09-674-379a-14
Perfect score: 2407
Sequence: 1 OCTGNFDDLRDROSGCCLDIDE.....INFRGSSVIRLRIVGQYPR 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneeqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneeqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneeqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneeqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneeqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneeqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneeqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneeqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneeqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneeqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneeqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneeqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneeqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneeqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneeqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneeqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneeqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneeqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneeqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneeqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneeqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneeqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	100.0	448	19	AAW79739 Human EGF protein
2	2407	100.0	448	20	AAW95709 Homo sapiens fetal
3	2407	100.0	448	20	AAW94281 Human extracellular
4	2407	100.0	448	21	AAV57058 Amino acid sequenc
5	2407	100.0	448	21	AAV54989 Full length human
6	2407	100.0	448	22	AAW93573 Human polypeptide
7	2407	100.0	448	23	AAU75494 Human extracellular
8	2401	99.8	448	20	AAV08063 Human EGF-like hom
9	2401	99.8	448	22	AAU29227 Human PRO polypept
10	2401	99.8	448	22	AAW31183 Amino acid sequenc

11	2302	95.6	423	21	AAV56751
12	2302	95.6	423	21	AAV56753
13	2302	95.6	448	21	AAV56750
14	2302	95.6	448	21	AAV54990
15	2302	95.6	461	21	AAV56752
16	2302	95.6	461	21	AAV54991
17	2230	92.6	392	18	AAW31705
18	1827	75.9	335	21	AAV76008
19	1827	75.9	335	22	AAW55947
20	1827	75.9	335	22	AAW72147
21	1276.5	53.0	443	18	AAW32110
22	1276.5	53.0	443	20	AAV16587
23	1276.5	53.0	443	21	AAW33418
24	1276.5	53.0	443	21	AAW84707
25	1276.5	53.0	443	21	AAV55850
26	1276.5	53.0	443	22	AAU12330
27	1276.5	53.0	443	21	AAV84706
28	1272.5	52.9	443	21	AAV84706
29	1270.5	52.8	443	22	AAW92533
30	1237.5	51.4	433	21	AAW58353
31	1106	45.9	201	22	AAW93220
32	1093.5	45.4	493	22	AAW72892
33	1093.5	45.4	493	22	AAW84707
34	992.5	41.2	295	21	AAV84705
35	989.5	41.1	295	21	AAV84708
36	963.5	40.0	350	20	AAV08066
37	963.5	40.0	387	22	AAW68188
38	912.5	37.9	274	21	AAV76081
39	912.5	37.9	274	22	AAW56020
40	912.5	37.9	274	23	AAW72220
41	811.5	33.7	413	22	AAW80391
42	811.5	33.7	413	23	AAW65300
43	811.5	33.7	451	22	AAW80440
44	735.5	30.6	683	12	AAW11150
45	735.5	30.6	683	18	AAW27600

ALIGNMENTS

RESULT 1	AAW79739	standard; Protein; 448 AA.
ID	AAW79739	
XX	AAW79739;	
AC	25-JAN-1999	(first entry)
XX		
DT		
XX		
DE		
XX		
KW	Human EGF protein.	
KW	Extracellular/epidermal growth factor-like protein; EGF; human; liver;	
KW	vascular smooth muscle cell proliferation; neurology; pathology; AIDS;	
KW	wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;	
KW	Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasma;	
KW	epidermal cell; cancer; psoriasis; detection.	
XX		
OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
FT	1..25	
FT	/label= signal	
FT	26..448	
FT	/label= EGF	
FT	/note= "Extracellular/epidermal growth factor-like protein"	
FT	112..153	
FT	/label= EGF-1	
FT	154..190	
FT	/label= EGF-2	
FT	191..230	
FT	/label= EGF-3	
FT	231..271	
FT	Region	

Smooth muscle prol
Smooth muscle prol
Smooth muscle prol
Full length mouse
Smooth muscle prol
Full length mouse
Human extracellular
Rat EGF extracellu
Skin cell protein,
Rat protein isolat
Human extracellular
Extracellular prot
Human PRO226 prote
A human p53 mutat
Human S1-5 ECPM-11
Human PRO226 POLYP
Human PRO226 polyp
Amino acid sequenc
Human protein sequ
Lung cancer associ
Human polypeptide,
Human EFEMP1. Hom
Human extracellular
The C-terminal of
C-terminal of p53
Human EGF-like pro
Extracellular prot
Human EGF extracel
Skin cell protein,
Human protein isol
Secreted protein e
Human albumin fusi
Gene #21 associate
Fibulin C. Homo s
Human fibulin type

Sat Jul 5 15:31:26 2003

us-09-674-379a-14.rai

Page 9

```

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (512) 418-3000
? TELEFAX: (713) 789-2679
? TELER: 79-0924
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1251 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
PCT-US95-02251-3

```

Query Match	18.28;	Score 438;	DB 5;	Length 1251;
Best Local Similarity	29.28;	Pred. No. 1.2e-26;		
Matches 109;	Conservative 36;	Mismatches 114;	Indels 114;	Gaps 15

```

0Y      2 CTNGFDLDRNOSQCLDIDECRTIPEAC--RGDMKCVNQNQGYLCIPRTNPVYRGPYSNPYS 60
Db      580 CNAQYKSHPHQHYRYCDVYNCECA--EPCCGKCKICNTGSSYNC-----620
0Y      61 PEYSGPYPAABPLSAPNYPITISRLICRFYQMD---ESNOCUDVDECATDSHQCNPTQ 117
Db      621 -----HCNREGRLRVGAGGRSCVDLINECA-KPHLCGDGG 653
0Y      118 ICINTEGGYTSCSTDGYWLLSEQ---CLDIDECR-YGYCQ-ICANVGSYSYC-TCNNGF 171
Db      654 FCINPFGHYKNCYRGYRLKMSRRPICDIDECRDPSTCPDGCKENKPGSPKCIACOPGY 713
0Y      172 TLNEGRCSCQDVNECATENPC-----VQT-----195
Db      714 R-SGGGACRBYNNECSSECTPCSPGWCEKLPGSRYCTCAQGIARTTGRSLCTDVIDDCEAGK 772
0Y      196 -----CVNTGSFYRCRCDPGYELEEDGVHRSMDDECSFSEFLCOHBCVNOPGTYPCSCP 249
Db      773 VCDGICICNTGSPQCCOLCSGLSHLSRORSREDIDEDCPAACIGCGCICINTNGSYRCLCP 832
0Y      250 PGYILLDNRSQDINECEGHNHTCNLQOTCYCNLOGGFKCIDPIRCEPYLRKISDNKCMC 309
Db      833 LGHRLVYGRKCKKDIDECSDQDPGLC-LPHACENLQGSYCV---CDEGGLTLPQDQH---884
0Y      310 PAENPGR--DQP 320
Db      885 -----GCEEVEOP 892

```

Search completed: July 3, 2003, 18:26:54
Job time : 13.1125 secs


```

Db      674 ECLNBS-TC-PDEQVNSPGSYOCVPTGFRGMNGCLDVLDECLBPNVCANGDCSNLEG 731
QY      162 SYSTCNPGFTLNEDGRSCODVNECATENPCVQ----- 194
Db      732 SYMCSHKYTRTPDHRKRDIDECQGNLCVNGCKNTEGSRCTCGGQYOLSAKDDC 791
QY      195 -----TCVNTYGSFTICRDPGYELEBDGVHSCSDMDECSFSEFICQ-HCV 238
Db      792 EDIDECOHNLCAHGCGRNTEGSPQCVCDOGYRASLGDHCEIDINCELEKSYCQGRDCT 851
QY      239 NOPGTFYSCPPGYILLDNRSQDINEGHNHTCNLDQTCYNLGGFKCI----- 290
Db      852 NTAGSYDCTCPGFG-QLDNKTCTQDINECEHFG-LCGPQGECLNTEGSEFHCVCQGGFSTIS 909
QY      291 -DPIICEEPYLRIS-----DN-----RCMC-----PAENPGCRDQPTILYRDM 328
Db      910 ADGRCEIDIDECVNTVCDSHGFCNNTAGSFRLCTGQGFQARQDGGCVD-----VNEC 963
QY      329 DVSIG 333
Db      964 ELISG 968

```

```

RESULT 12
US-08-897-443-1
; Sequence 1, Application US/08897443
; Patent No. 5981263
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Kaaser, Matthew
; TITLE OF INVENTION: HUMAN MATRILIN-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,443
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0348 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRMOT02
; CLONE: 681719
; US-08-897-443-1

```

```

Query Match      18.9% Score 455; DB 2; Length 638;
Best Local Similarity 28.1%; Pred. No. 2.7e-28;

```

```

Matches 92; Conservative 59; Mismatches 112; Indels 64; Gaps 11;
QY      1 QCTNGFPLDRSGGCLDIDECRTIPENACRGDMCMYNGYLCITRTPYRGPSPNPS 60
Db      304 QCTSYALAEGRKCAVADVTCASRNHGCHE--CVNADGSLC----- 344
QY      61 TPVSGPYPAAPPLSAPNYPPIISRPLICRFQYOMD-ESNOCVDVDECATDSHOCNPQIC 119
Db      345 -----QCHEGFALNPDKKITKIDYCASSNHGCO--HEC 376
QY      120 INTERGYTSGCTGSYMLLEGQ--CLDIDECRYG--COQLCANPGYSTCNPGETLINE 175
Db      377 VNTDDSTSCHLKGFETLNPDKTCRRINYLALNPGCEHECVNMEESYCRHGYTLDP 436
QY      176 DGRSCODVNECA--TENPCVQTCVNTYGSFTICRDPGYELEBDGVHSCSDMDECSFSEFICQ 234
Db      437 NGKTCSRVDHCAQADHGCEDCLNTEDESPFQCCEGFLINDLTCSRVDDICLLSDHGC 496
QY      235 HECVNOPTGYFCSCPFGYILLDNRSQDINEGHNHTCNLDQTCYNLGGFKCIPIR 294
Db      497 YSCVNMDSRFACQCPFGHYLRSDGKTCARLIDSCALGDHGC--EHSVSSSDSFFVC---Q 550
QY      295 CEEPYLRISDNR-C---MCPAENPGC 316
Db      551 CEEGYTLREDGKTCRRKDVCAIDHGC 577

```

```

RESULT 13
US-08-479-722B-2
; Sequence 2, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yan, Wushan
; TITLE OF INVENTION: LATENT TGF $\beta$  BINDING PROTEIN (LTBP)
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.
; REGISTRATION NUMBER: 39,458
; REFERENCE/DOCKET NUMBER: 4100.000500/FUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 934-7000
; TELEFAX: (713) 934-7011
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1833 amino acids
; TYPE: amino acid

```

```

; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/84,072
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; FAX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 458228
; US-09-212-168-5

Query Match 40.0%; Score 963.5; DB 4; Length 387;
Best Local Similarity 48.3%; Pred. No. 1.1e-68;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

QY 71 APPLSAPNTPTISRLPRLICRGYOMDESNOQCVDECATDSHCNPTQICINTEGGTSC 130
DB 38 ADPQRIPSNP--SHRIQACAGIESEHNVQDIDECTAGTNCADQYICINLRGSFACOC 95
QY 131 TGGYLLBEGCGLDIDECRY-GYCOQLCANYPGYSCTCNPGLTINEDGRSCODVNECATE 189
DB 96 PGYQCRGECQVIDECTIPPYCHQRCVNTPGSFYCCSPFQOLANNYTCVDINECDAS 155
QY 190 NPCVQCVMTYSGFICRCDDPGYELLEDGVCNCSMDDECSFSLQHECVNPGTYFCGSCP 249
DB 156 NCAQCCYNILGSLFTCCQNGYELSDRLNCEDIDERTSSYLQYQCVNEPGKFSMCP 215
QY 250 PGYILLDRNSCODINECEHRNHTCNLQOTCYNLQGGFKCIDPRLCEBPYLRISDNRCM 309
DB 216 QGYQYV-RSTTCODINECEFTNE-CRDEDMCNHGGFRCPRLNCCQDPYILTPENRCYC 273
QY 310 PAENPCRODPTILYRDMDVSGRSVPADIFQMAATRYPGAYYIFQIKSGNEGREFFM 369
DB 274 PVSNAACRELPOSIVYKYSIRSDRSVPSDIFQIQAATYVANTINTFRKSGNENGEFYL 333
QY 370 ROTGPTSATLVMTRPIKGRPREIQLDLEMTVNTYINFRGSSVIRLRIYVQYRF 423
DB 334 ROTSPVSAVLVKSLSGPREHIVDLEMLTVSSIGTERTSVLRLLTIIVGPFSS 387

```

```

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 336
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Human
; US-09-188-930-336

Query Match 37.9%; Score 912.5; DB 4; Length 274;
Best Local Similarity 56.0%; Pred. No. 8e-65;
Matches 154; Conservative 47; Mismatches 73; Indels 1; Gaps 1;

QY 149 YGYCOQLCANYPGYSCTCNPGLTINEDGRSCODVNECATENPCVQCVNTYGSFICRD 208
DB 1 YRYCQHRVNLPGSFRCOCCEBFGOLGPNRNSCVDNECDMGAPCBQRCFNSYGTFLCRCH 60
QY 209 PGYELLEDGYHCSMDCSFSEFLCQHECVNPGTYFCSCPPGYILLDRNSCODINECE 268
DB 61 QGYELHRDGFSCSDIDECYSYSLCYQRCVNEPGFSCHPGYQLL-ATRLCQDIDEC 119
QY 269 HRHTCNLQOTCYNLQGGFKCIDPRLCEBPYLRISDNRCMCPAENPGCRDQPTILYRDM 328
DB 120 SGAHQCEAQTQVAFHGGYRVDINRCVEPIQYQSENCLCPASNPICREPSIYHRM 179
QY 329 DVSGRSVPADIFQMAATRYPGAYYIFQIKSGNEGREFFYRQTPISATLVMTRPIKCP 388
DB 180 TITSERSVPADVFEIQATSVYPGAYYNAFQIRAGNSQGFYIROLNNSAMVLARVYGP 239
QY 389 RETOLDLEMTVNTYINFRGSSVIRLRIYVQYRF 423
DB 240 REYVLDEMTVNTSLMSTRASSVLRLLTVGAYTF 274

RESULT 11
5177197-30
; PATENT NO. 5177197
; APPLICANT: KANZAKI, TETSUO; OLOESSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,
; LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 30
; LENGTH: 1394
; 5177197-30

Query Match 22.4%; Score 539; DB 6; Length 1394;
Best Local Similarity 29.6%; Pred. No. 1.5e-34;
Matches 126; Conservative 51; Mismatches 126; Indels 122; Gaps 17;

QY 2 CTGPFDRSGGCLDIDECRTIPEACRGDMCNQNGCYLCIPRTNYYRPPYSPYST 61
DB 573 CYEGYRSEQORCCVDIDECTQYQHL-C-SQRCENTESSFLCI----- 614
QY 62 PYSGPYPAAPPLSAPNTPTISRLP-----CRFGYOMDESNOQCVVD 104
DB 615 -CPAGFMASEGRNCIDVDECLRPDYCGEGHCYNTVGAFRCEYCDSGYRMYRGCEID 673
QY 105 ECATDGHQCNPTQICINTEGGTTC-SCTDGYLLBEGCGLDIDEC-RYGYCOQ-LCANYPG 161

```

```

: CITY: Rockville
: STATE: MD
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/833,963C
: FILING DATE: 11-APR-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US96/05033
: FILING DATE: 10-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PF258
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 301-309-8504
: TELEFAX: 301-309-8439
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 387 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-833-963C-9

```

```

Query Match      40.0%; Score 963.5; DB 2; Length 387;
Best Local Similarity 48.3%; Pred. No. 1,1e-68;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

: 71 APPLSANPYTISPLICRGYOMDESNOCVDVDECATDSHQCNPOTICINTEGGYTCSC 130
: 38 ADQRIPSNP--SHRICAGAGYEQSEHNVCODIDECTAGHNCRADQVCINLRGSPACQC 95
: 131 TDGYWLLGCGCLDIDECRY-GYCOOLCANYPGSYSTCNGFTLNEDGRSCODVNECATE 189
: 96 PPGYQKRGECQVVIDECTIPPYCHQRCVNPFGSFYCCSGFOLAANNYTCVADINECDAS 155
: 190 NPCVQTCVNTYGSFICRCDPGVELEEDGVHSCDMECSSEFLCQHECVNQPGTYFCSCP 249
: 156 NQCAQCYNILGSLFICQNGYELSSDRNCEDIDECRTSSYLCOYQCVNEPKGFCSCMP 215
: 250 PGYILLDNRSQDINECEHNRHTCNLQOTCYNLOGGFCIDPICEEPLRLISDRNCMC 309
: 216 OGQYVV-RSRTCODINECETTNE-CREDEMCWNYHGFRCYRPNPCQDPYILLPENRCVC 273
: 310 PAENPGCRPOPTIILYRDMDVYSGRSVPADIFOMATTRYPGAYVYFQIKSGNEGREFYA 369
: 274 PVSNAACRLPDSIYKTKYSIRSDRSVPDSIDFOIQTATTIYANTINFRKSGNEGEFYL 333
: 370 ROTGPISATLVMTRPKGRPREIOLDLEMTVTVINFRGSSVIRLRIYSOYRF 423
: 334 ROTSPVSAMLVYKSLSGPREHIYDLEMTVTSIGTFRISVLRLLTIIVGPSPF 387

```

```

: RESULT 8
: US-08-980-514-3
: Sequence 3, Application US/08980514
: Patent No. 6004753
: GENERAL INFORMATION:
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Shah, Purvi
: TITLE OF INVENTION: HUMAN SI-5-DCMP-LIKE PROT
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Dr.
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/980,514
: FILING DATE: Filed Herewith
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0436 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-845-0555
: TELEFAX: 650-845-4166
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 387 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 458228
: US-08-980-514-3

```

```

Query Match      40.0%; Score 963.5; DB 3; Length 387;
Best Local Similarity 48.3%; Pred. No. 1,1e-68;
Matches 171; Conservative 55; Mismatches 123; Indels 5; Gaps 4;

: 71 APPLSANPYTISPLICRGYOMDESNOCVDVDECATDSHQCNPOTICINTEGGYTCSC 130
: 38 ADQRIPSNP--SHRICAGAGYEQSEHNVCODIDECTAGHNCRADQVCINLRGSPACQC 95
: 131 TDGYWLLGCGCLDIDECRY-GYCOOLCANYPGSYSTCNGFTLNEDGRSCODVNECATE 189
: 96 PPGYQKRGECQVVIDECTIPPYCHQRCVNPFGSFYCCSGFOLAANNYTCVADINECDAS 155
: 190 NPCVQTCVNTYGSFICRCDPGVELEEDGVHSCDMECSSEFLCQHECVNQPGTYFCSCP 249
: 156 NQCAQCYNILGSLFICQNGYELSSDRNCEDIDECRTSSYLCOYQCVNEPKGFCSCMP 215
: 250 PGYILLDNRSQDINECEHNRHTCNLQOTCYNLOGGFCIDPICEEPLRLISDRNCMC 309
: 216 OGQYVV-RSRTCODINECETTNE-CREDEMCWNYHGFRCYRPNPCQDPYILLPENRCVC 273
: 310 PAENPGCRPOPTIILYRDMDVYSGRSVPADIFOMATTRYPGAYVYFQIKSGNEGREFYA 369
: 274 PVSNAACRLPDSIYKTKYSIRSDRSVPDSIDFOIQTATTIYANTINFRKSGNEGEFYL 333
: 370 ROTGPISATLVMTRPKGRPREIOLDLEMTVTVINFRGSSVIRLRIYSOYRF 423
: 334 ROTSPVSAMLVYKSLSGPREHIYDLEMTVTSIGTFRISVLRLLTIIVGPSPF 387

```

```

: RESULT 9
: US-09-212-168-5
: Sequence 5, Application US/09212168
: Patent No. 6303765
: GENERAL INFORMATION:
: APPLICANT: Bandman, Olga
: APPLICANT: Corley, Neil C.
: APPLICANT: Guegler, Karl J.
: TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

```


Matches	423;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	QCTNGFGLDRSGGCLDIDECRTIP	EACRGDM	CYVNGGYLCIP	RNPVYRG	PSN	60		
Db	26	QCTNGFGLDRSGGCLDIDECRTIP	EACRGDM	CYVNGGYLCIP	RNPVYRG	PSN	85		
Qy	61	TPYSGPYAAAPPLSAPNYP	PTISRLP	ICRFGYOM	ESNOCV	VDVECATDS	HOCNPT	120	
Db	86	TPYSGPYAAAPPLSAPNYP	PTISRLP	ICRFGYOM	ESNOCV	VDVECATDS	HOCNPT	145	
Qy	121	NTEGGYTCSTDGYWLL	EGGCLDIDECRY	GYCQQLCANV	PGSYSTC	NPGEFTL	NEDGRSC	180	
Db	146	NTEGGYTCSTDGYWLL	EGGCLDIDECRY	GYCQQLCANV	PGSYSTC	NPGEFTL	NEDGRSC	205	
Qy	181	ODVNECATENNCVQTCV	NTVTSFICRCP	DGYELEEDGYH	CSMDCSSE	FLCQHE	CYVQ	240	
Db	206	ODVNECATENNCVQTCV	NTVTSFICRCP	DGYELEEDGYH	CSMDCSSE	FLCQHE	CYVQ	265	
Qy	241	PGTFCSGCPGYILL	DNRSQCQDINE	CHRNHTCM	LAOTCYNL	OGFACIDP	IRCEEPYL	300	
Db	266	PGTFCSGCPGYILL	DNRSQCQDINE	CHRNHTCM	LAOTCYNL	OGFACIDP	IRCEEPYL	325	
Qy	301	RISDNRCMCPEAN	EGCRDQPTILY	RDMDVY	SGRSVPADIF	QOMATTRY	PGAYYIFQIKS	360	
Db	326	RISDNRCMCPEAN	EGCRDQPTILY	RDMDVY	SGRSVPADIF	QOMATTRY	PGAYYIFQIKS	385	
Qy	361	GNEGREFMRQTGP	LSATLVMT	TRPIKGR	REIQLD	LEMTVNTV	INFRSSVIRL	420	
Db	386	GNEGREFMRQTGP	LSATLVMT	TRPIKGR	REIQLD	LEMTVNTV	INFRSSVIRL	445	
Qy	421	YPF	423						
Db	446	YPF	448						

RESULT 2
US-09-212-168-1
Sequence 1, Application US/09212168
Patent No. 6303765
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/212,168
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/884,072
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 448 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: CORNMO101
: CLONE: 45517
:
: US-09-212-168-1
:
: Query Match 100.0%; Score 2407; DB 4; Length 448;
: Best Local Similarity 100.0%; Pred. No. 1,46-182;
: Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 1 OCTNGFELDDRSSGCLDIDECRTTPEACRGDMCYNONGGYLCIPRTNPVYRGPSNPYS 60
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 26 QCTNGFELDDRSSGCLDIDECRTTPEACRGDMCYNONGGYLCIPRTNPVYRGPSNPYS 85
: Oy 61 TPYSGPYAAPAPPLSAPNPPTISRPPLICRFGYOMDESNCVYVDECATDSHOCNPTQICI 120
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 86 TPYSGPYAAPAPPLSAPNPPTISRLICRFGYOMDESNCVYVDECATDSHOCNPTQICI 145
: Oy 121 NTEGGYTCSCTDGYWLLLEGQCLDIDECRYGYCOQLCANVPSSYSCTCNAGFTLNEGGRSC 180
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 146 NTEGGYTCSCTDGYWLLLEGQCLDIDECRYGYCOQLCANVPSSYSCTCNAGFTLNEGGRSC 205
: Oy 181 QDVNECATENPCVQTCVNTGTSFICRCRCPGYELEEDGYHCSMDSCSFSEFLCOHECVNQ 240
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 206 QDVNECATENPCVQTCVNTGTSFICRCRCPGYELEEDGYHCSMDSCSFSEFLCOHECVNQ 265
: Oy 241 PGTFYSCPPYILLDDNRSCODINECHRNHTCNLDQTCYNLDGGFKCIDPIRCEEPYL 300
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 266 PGTFYSCPPYILLDDNRSCODINECHRNHTCNLDQTCYNLDGGFKCIDPIRCEEPYL 325
: Oy 301 RISDNRCPCAPENPCRCRQPPFTILYRDMDVYSGRVPADIFQMATTRTPGAYYIFQIKS 360
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 326 RISDNRCPCAPENPCRCRQPPFTILYRDMDVYSGRVPADIFQMATTRTPGAYYIFQIKS 385
: Oy 361 GNEGRFPYMRQTPISATLVMTRPDKPGREIQLDLEMITVNVINFRSSVIRLRIVYQS 420
: Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: 386 GNEGRFPYMRQTPISATLVMTRPDKPGREIQLDLEMITVNVINFRSSVIRLRIVYQS 445
:
: Oy 421 YPF 423
: | | |
: Db 446 YPF 448
:
: RESULT 3
: US-09-188-930-186
: Sequence 186, Application US/09188930A
: Patent No. 6150502
: GENERAL INFORMATION:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
: APPLICANT: Murison, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 186
: LENGTH: 337
: TYPE: PRT
: ORGANISM: Rat
: FEATURE:
: NAME/KEY: UNSURE
: LOCATION: (14)...(124)
: NAME/KEY: UNSURE
: LOCATION: (135)...(135)

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:54:52 ; Search time 13.1125 Seconds
(Without alignments)
949.162 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFDLRQSGQCLDIDE.....INFRGSSVIRLRIVSQYF 423

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2407	100.0	448	2	US-08-884-072-1
2	2407	100.0	448	4	US-09-212-168-1
3	1791	74.4	337	4	US-09-188-930-186
4	1276.5	53.0	443	2	US-08-833-963C-2
5	1276.5	53.0	443	3	US-08-980-514-1
6	963.5	40.0	387	2	US-08-884-072-5
7	963.5	40.0	387	2	US-08-833-963C-9
8	963.5	40.0	387	3	US-08-980-514-3
9	963.5	40.0	387	4	US-09-212-168-5
10	912.5	37.9	274	4	US-09-188-930-336
11	539	22.4	1394	6	517197-30
12	455	18.9	638	2	US-08-897-443-1
13	453	18.8	1833	3	US-08-479-722B-2
14	453	18.8	1833	3	US-08-479-722B-2
15	438	18.2	1251	5	PCT-US95-02251-18
16	438	18.2	1252	1	US-08-199-780-3
17	438	18.2	1252	2	US-08-316-650-3
18	438	18.2	1253	3	US-08-479-722B-4
19	431	17.9	956	2	US-08-897-443-3
20	392	16.3	886	3	US-09-110-116-3
21	372	15.5	810	2	US-08-820-170A-34
22	372	15.5	810	3	US-09-055-699-34
23	372	15.5	810	4	US-09-273-565-34
24	372	15.5	810	4	US-09-565-538-34
25	372	15.5	810	4	US-09-661-468-34
26	369	15.3	2471	1	US-08-185-432-16
27	369	15.3	2471	1	US-08-083-590A-19

28	369	15.3	2471	3	US-08-532-384-19	Sequence 19, Appl
29	369	15.3	2471	4	US-08-899-232-1	Sequence 1, Appl
30	361.5	15.0	816	2	US-08-820-170A-37	Sequence 37, Appl
31	361.5	15.0	816	4	US-09-055-699-37	Sequence 37, Appl
32	361.5	15.0	816	4	US-09-273-565-37	Sequence 37, Appl
33	361.5	15.0	816	4	US-09-565-538-37	Sequence 37, Appl
34	361.5	15.0	816	4	US-09-661-468-37	Sequence 37, Appl
35	354.5	14.7	1964	4	US-09-467-997-1	Sequence 1, Appl
36	351.5	14.6	2703	1	US-08-185-432-19	Sequence 19, Appl
37	351.5	14.6	2703	4	US-08-899-232-4	Sequence 4, Appl
38	350.5	14.6	652	2	US-08-751-305-2	Sequence 2, Appl
39	345.5	14.4	2556	1	US-08-083-590A-20	Sequence 20, Appl
40	345.5	14.4	2556	3	US-08-532-384-20	Sequence 20, Appl
41	342.5	14.2	2556	1	US-08-185-432-17	Sequence 17, Appl
42	342.5	14.2	2556	4	US-08-899-232-2	Sequence 3, Appl
43	340.5	14.1	673	1	US-08-282-141-3	Sequence 3, Appl
44	340.5	14.1	673	1	US-08-435-434-1	Sequence 1, Appl
45	340.5	14.1	673	1	US-08-435-436-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-884-072-1
Sequence 1, Application US/08884072
Patent No. 5872234
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OR INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOT01
CLONE: 45517
US-08-884-072-1
Query Match 100.0%; Score 2407; DB 2; Length 448;
Best Local Similarity 100.0%; Pred. No. 1.4e-182;


```
QY 1 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPNPYS 60
Db 26 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPNPYS 85
QY 61 TPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDENOCVDVDECATDSHOCNPTQICI 120
Db 86 TPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDENOCVDVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTDGWVLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEGDRSC 180
Db 146 NTEGGYTCSTDGWVLEGGCLDIDECRYGYCQQLCANVPGSYSCTCNPGFTLNEGDRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYELEEDGYHCSMDDECSFSEFLCOHECVNO 240
Db 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYELEEDGYHCSMDDECSFSEFLCOHECVNO 265
QY 241 PGTYFGSCPFGYILLDDNRSCODINECEHRNHTCNLQOTCYNLGGGFKCIDPIRCEEPYL 300
Db 266 PGTYFGSCPFGYILLDDNRSCODINECEHRNHTCNLQOTCYNLGGGFKCIDPIRCEEPYL 325
QY 301 RISDNRCMCPAENBPCRDOPFTIILYRMDVYSGRSVPADIFOMOATTRYPGAYYIPIQIS 360
Db 326 RISDNRCMCPAENBPCRDOPFTIILYRMDVYSGRSVPADIFOMOATTRYPGAYYIPIQIS 385
QY 361 GNEGREFFYMRQTPISATLVMTPIKGPRIQLDLEMITVNTVINFRGSSVIRLRIYVSQ 420
Db 386 GNEGREFFYMRQTPISATLVMTPIKGPRIQLDLEMITVNTVINFRGSSVIRLRIYVSQ 445
QY 421 YPF 423
Db 446 YPF 448
```

Search completed: July 3, 2003, 18:26:16
Job time : 83.5316 secs

```

; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLROSQCLDIDECRTIPEACRGDMKCVNONGYLCTPRNPVYRGPSNYS 60
DB 26 OCTNGFDLROSQCLDIDECRTIPEACRGDMKCVNONGYLCTPRNPVYRGPSNYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRGYQWDESNOCVDDECATDSHCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPLICRGYQWDESNOCVDDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSDTGWYLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 180
DB 146 NTEGGYTCSDTGWYLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 240
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 265
QY 241 PGTYFCSPPGYILLDDNRSQDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 300
DB 266 PGTYFCSPPGYILLDDNRSQDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 325
QY 301 RISDNRCMCPAENPGCRDOPFTLLYRDMDVSGRSVPADIFOMATRRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDOPFTLLYRDMDVSGRSVPADIFOMATRRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVNTVNFSGSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVNTVNFSGSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 14
US-10-176-913-408
; Sequence 408, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
```

```

; ORGANISM: Homo Sapien
US-10-176-913-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLROSQCLDIDECRTIPEACRGDMKCVNONGYLCTPRNPVYRGPSNYS 60
DB 26 OCTNGFDLROSQCLDIDECRTIPEACRGDMKCVNONGYLCTPRNPVYRGPSNYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRGYQWDESNOCVDDECATDSHCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNPTISRPLICRGYQWDESNOCVDDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSDTGWYLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 180
DB 146 NTEGGYTCSDTGWYLLLEGQCLDIDECRYGCOQLCANVPGSYSCCTNPFTLNEDGRSC 205
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 240
DB 206 ODVNECATENPCVQTCVNTYGSFICRCDDPGYTELEEDGVHSDMECSFSEFLCOHECVNO 265
QY 241 PGTYFCSPPGYILLDDNRSQDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 300
DB 266 PGTYFCSPPGYILLDDNRSQDINECEHNNHTCNLOQTCYNLQGFRCIDPICEEPLY 325
QY 301 RISDNRCMCPAENPGCRDOPFTLLYRDMDVSGRSVPADIFOMATRRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDOPFTLLYRDMDVSGRSVPADIFOMATRRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVNTVNFSGSVIRLRIVSQ 420
DB 386 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVNTVNFSGSVIRLRIVSQ 445
QY 421 YPF 423
DB 446 YPF 448

RESULT 15
US-10-180-552-408
; Sequence 408, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C13
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-408

Query Match
Best Local Similarity 99.8%; Score 2401; DB 9; Length 448;
Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT APPLICATION NUMBER: US/10/175,752
; PRIORITY FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
; 10-175-752-408

```

```

Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPSNPYS 60
DB 26 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPSNPYS 85
QY 61 TPYSGPYPAAPPLSAPNTPTISRPILCRFGYOMDESNOCVYDDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNTPTISRPILCRFGYOMDESNOCVYDDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPFGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPFGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYELFEDDVHCSMDCESEFICQHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYELFEDDVHCSMDCESEFICQHECVNQ 265
QY 241 PGTYFSCPPGYILLDNRSQCQDINECEHNRHTCNLOQTCYNLOGGFKCIDPIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDNRSQCQDINECEHNRHTCNLOQTCYNLOGGFKCIDPIRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
QY 361 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRITYSQ 420
DB 386 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

```

RESULT 12
US-10-176-482-408
; Sequence 408, Application US/10176482
; Publication No. US20030022296A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

```

```

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; PRIORITY FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
; 10-176-482-408

```

```

Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2,3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPSNPYS 60
DB 26 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPVYRGPSNPYS 85
QY 61 TPYSGPYPAAPPLSAPNTPTISRPILCRFGYOMDESNOCVYDDECATDSHOCNPTQICI 120
DB 86 TPYSGPYPAAPPLSAPNTPTISRPILCRFGYOMDESNOCVYDDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPFGFTLNEDGRSC 180
DB 146 NTEGGYTCSTGTYWLLLEGQCLDIDECRYGYCOQLCANVPGSYSCTCNPFGFTLNEDGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYELFEDDVHCSMDCESEFICQHECVNQ 240
DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYELFEDDVHCSMDCESEFICQHECVNQ 265
QY 241 PGTYFSCPPGYILLDNRSQCQDINECEHNRHTCNLOQTCYNLOGGFKCIDPIRCEEPYL 300
DB 266 PGTYFSCPPGYILLDNRSQCQDINECEHNRHTCNLOQTCYNLOGGFKCIDPIRCEEPYL 325
QY 301 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 RISDNRCMCPAENPGCRDQPTIILYRDMDVVSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
QY 361 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRITYSQ 420
DB 386 GNGREFFYMRQTGPISATLVMTPIKGPRTIOLDLEMTVNTVINFRGSSVIRLRITYSQ 445
QY 421 YPF 423
DB 446 YPF 448

```

```

RESULT 13
US-10-176-757-408
; Sequence 408, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86

```

QY 361 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 420
| | | | |
Db 386 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 445

QY 421 YPF 423
| | |
Db 446 YPF 448

RESULT 9

US-10-173-706-408
; Sequence 408, Application US/10173706
; Publication No. US20030022293A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
PRIORITY FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMACVNOGNGYLCPRTNPPYRGYSNPYS 60
| | | | |
Db 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMACVNOGNGYLCPRTNPPYRGYSNPYS 85
| | | | |
Y 61 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMDESNOQVDVDECATDHSQCNPOTICI 120
| | | | |
Db 86 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMDESNOQVDVDECATDHSQCNPOTICI 145
| | | | |
QY 121 NTEGGYTSCDTGYWLLBEGOCCLDIDECRYGCOQLCANVGSYCTCNPGFTLNEDGRSC 180
| | | | |
Db 146 NTEGGYTSCDTGYWLLBEGOCCLDIDECRYGCOQLCANVGSYCTCNPGFTLNEDGRSC 205
| | | | |
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 240
| | | | |
Db 206 ODVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 265
| | | | |
QY 241 PGTFECSGPPGYILLDDNRSODINCEHRNHTCNLOQTCYNLOGGKCIDPICEBPYL 300
| | | | |
Db 266 PGTFECSGPPGYILLDDNRSODINCEHRNHTCNLOQTCYNLOGGKCIDPICEBPYL 325
| | | | |
QY 301 RISDNRCMCPAENPCGRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 360
| | | | |
Db 326 RISDNRCMCPAENPCGRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 385
| | | | |
QY 361 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 420
| | | | |
Db 386 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 445
| | | | |
QY 421 YPF 423
| | |
Db 446 YPF 448

RESULT 10

US-10-175-738-408
; Sequence 408, Application US/10175738
; Publication No. US20030022294A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
PRIORITY FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;

Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMACVNOGNGYLCPRTNPPYRGYSNPYS 60
| | | | |
Db 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMACVNOGNGYLCPRTNPPYRGYSNPYS 85
| | | | |
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMDESNOQVDVDECATDHSQCNPOTICI 120
| | | | |
Db 86 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMDESNOQVDVDECATDHSQCNPOTICI 145
| | | | |
QY 121 NTEGGYTSCDTGYWLLBEGOCCLDIDECRYGCOQLCANVGSYCTCNPGFTLNEDGRSC 180
| | | | |
Db 146 NTEGGYTSCDTGYWLLBEGOCCLDIDECRYGCOQLCANVGSYCTCNPGFTLNEDGRSC 205
| | | | |
QY 181 ODVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 240
| | | | |
Db 206 ODVNECATENPCVQTCVNTYGSFICRCDPGYLEEDGVHCSMDDECSFSEFLCOHECVNQ 265
| | | | |
QY 241 PGTFECSGPPGYILLDDNRSODINCEHRNHTCNLOQTCYNLOGGKCIDPICEBPYL 300
| | | | |
Db 266 PGTFECSGPPGYILLDDNRSODINCEHRNHTCNLOQTCYNLOGGKCIDPICEBPYL 325
| | | | |
QY 301 RISDNRCMCPAENPCGRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 360
| | | | |
Db 326 RISDNRCMCPAENPCGRDQPFITLYRDMDVSGRSVPADIFQMATTRYGAYYIFQIKS 385
| | | | |
QY 361 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 420
| | | | |
Db 386 GNEGREFYMRQTGPISATLVMTRPRIKGPRIQDLEMTVNTVINFGSSVIRLRIVSQ 445
| | | | |
QY 421 YPF 423
| | |
Db 446 YPF 448

RESULT 11

US-10-175-752-408
; Sequence 408, Application US/10175752
; Publication No. US20030022295A1

GENERAL INFORMATION:

Db 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPGSYSTCNPFGTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 240
Db 206 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 300
Db 266 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 360
Db 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 385
QY 361 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVTVINFRSSVIRLRITYSO 420
Db 386 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVTVINFRSSVIRLRITYSO 445
QY 421 YPF 423
Db 446 YPF 448

RESULT 7

US-10-176-758-408
Sequence 408, Application US/10176758
Publication No. US20030008353A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-758-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPPVYRGPSNPS 60
Db 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPPVYRGPSNPS 85
QY 61 TPYSGPYPAAPLAPLADNPYPTISRPLICRFGYQMDENOCVADVDECATDSHOCNPTQICI 120
Db 86 TPYSGPYPAAPLAPLADNPYPTISRPLICRFGYQMDENOCVADVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPGSYSTCNPFGTLNEDGRSC 180
Db 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPGSYSTCNPFGTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 240
Db 206 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 300

Db 266 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 360
Db 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 385
QY 361 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVTVINFRSSVIRLRITYSO 420
Db 386 GNEGREFYMRQTPISATLWTRPIKPREIQLDLEMTVTVINFRSSVIRLRITYSO 445
QY 421 YPF 423
Db 446 YPF 448

RESULT 8

US-10-175-737-408
Sequence 408, Application US/10175737
Publication No. US20030013153A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C150
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 408
LENGTH: 448
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-408

Query Match 99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPPVYRGPSNPS 60
Db 26 OCTNGFDLDRSGQCLDIDECRTIPEACRGDMCMVONGGYLCIPRTNPPVYRGPSNPS 85
QY 61 TPYSGPYPAAPLAPLADNPYPTISRPLICRFGYQMDENOCVADVDECATDSHOCNPTQICI 120
Db 86 TPYSGPYPAAPLAPLADNPYPTISRPLICRFGYQMDENOCVADVDECATDSHOCNPTQICI 145
QY 121 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPGSYSTCNPFGTLNEDGRSC 180
Db 146 NTEGGYTCSTGDTGWLLEGGCLDIDECRYGCOOLCANVPGSYSTCNPFGTLNEDGRSC 205
QY 181 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 240
Db 206 QDVNECATENPCVQTCVNTGVSFICRCDPGYLEBEDVHCSMDCECSSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 300
Db 266 PGTYFSCPPGYILLDDNRSCODINECEHNRHTCNLOOTCYNLGGFCIDPIRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 360
Db 326 RISDNRCMCAENPGCRDOPFTILYRDMVYSGRSVPADIFOMATTTRPGAYYIFOLKS 385

```

; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403296
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/423844
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 09/522342
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: 09/548815
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 09/664610
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/665350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/767609
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/808689
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/866028
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/870574
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/886342
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: PCT/US98/14552
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US98/18824
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/19093
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: PCT/US98/19330
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: PCT/US98/19437
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/24855
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: PCT/US98/25108
; PRIOR FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: PCT/US98/25190
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: PCT/US99/20111
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNFPVRYGPNYS 60
Db 26 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNFPVRYGPNYS 85
```

```

QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESNOCVYDVECATDSHCNPOTICI 120
Db 86 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESNOCVYDVECATDSHCNPOTICI 145
QY 121 NTGSGYTCSTGDTWLEGGCLDIDECRYGYCOOLCANVPSYSCNPGFTLNEDGRSC 180
Db 146 NTGSGYTCSTGDTWLEGGCLDIDECRYGYCOOLCANVPSYSCNPGFTLNEDGRSC 205
QY 181 QDVNECATENPCVQOTCVNTGSPFCRCDDPCYELLEDVYHCSMDDECSFSEFLCOHECVNQ 240
Db 206 QDVNECATENPCVQOTCVNTGSPFCRCDDPCYELLEDVYHCSMDDECSFSEFLCOHECVNQ 265
QY 241 PGTYFSCPPGYLLDNRSCODINCEHNNHTCNLOOTCYNLOGGFKCIDPIRCEPPTL 300
Db 266 PGTYFSCPPGYLLDNRSCODINCEHNNHTCNLOOTCYNLOGGFKCIDPIRCEPPTL 325
QY 301 RSTDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSVPADIFQMATTIRYGAYYIFQIKS 360
Db 326 RSTDNRCMCPAENPGCRDQFTLLYRDMDVYSGRSVPADIFQMATTIRYGAYYIFQIKS 385
QY 361 GNEGREFYMRQTPISATLWMPRIKPREIOLDLEMTYNTYINRGSSVIRLRYVSQ 420
Db 386 GNEGREFYMRQTPISATLWMPRIKPREIOLDLEMTYNTYINRGSSVIRLRYVSQ 445
QY 421 YPF 423
Db 446 YPF 448

RESULT 6
US-10-174-590-408
; Sequence 408, Application US/10174590
; Publication No. US2003008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 408
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-408

Query Match          99.8%; Score 2401; DB 9; Length 448;
Best Local Similarity 99.8%; Pred. No. 2.3e-159;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNFPVRYGPNYS 60
Db 26 QCTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNFPVRYGPNYS 85
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESNOCVYDVECATDSHCNPOTICI 120
Db 86 TPYSGPYPAAPPLSAPNYPITISRLICRGYOMDESNOCVYDVECATDSHCNPOTICI 145
QY 121 NTGSGYTCSTGDTWLEGGCLDIDECRYGYCOOLCANVPSYSCNPGFTLNEDGRSC 180
Db 146 NTGSGYTCSTGDTWLEGGCLDIDECRYGYCOOLCANVPSYSCNPGFTLNEDGRSC 205
```

OY 361 GNEGRFYMRQTPISATLVMTFRPKPREIQLDLEMTVNTVINFRSSVIRLRIYVSO 420
DB 386 GNEGRFYMRQTPISATLVMTFRPKPREIQLDLEMTVNTVINFRSSVIRLRIYVSO 445

OY 421 YPF 423
DB 446 YPF 448

RESULT 5

US-10-066-500-15
Sequence 15, Application US/10066500
Patent No. US20020177165A1

GENERAL INFORMATION:

APPLICANT: Avi J. Ashkenazi
APPLICANT: Kevin P. Baker
APPLICANT: David A. Botstein
APPLICANT: Luc Desnoyers
APPLICANT: Dan L. Eaton
APPLICANT: Napoleone Ferrara
APPLICANT: Sherman Fong
APPLICANT: Wei-Olang Gao
APPLICANT: Hanspeter Gerber
APPLICANT: Mary E. Gerltsen
APPLICANT: Audrey Goddard
APPLICANT: Paul J. Godowski
APPLICANT: Austin L. Gurney
APPLICANT: Ivar J. Kjaevn
APPLICANT: Jennie P. Mather
APPLICANT: Mary A. Napier
APPLICANT: James Pan
APPLICANT: Nicholas F. Paoni
APPLICANT: Margaret Ann Roy
APPLICANT: Timothy A. Stewart
APPLICANT: Daniel Tumas
APPLICANT: Colin K. Watanabe
APPLICANT: P. Mickey Williams
APPLICANT: William I. Wood
APPLICANT: Zemin Zang

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS

FILE REFERENCE: F3130R1C7

CURRENT APPLICATION NUMBER: US/10/066, 500
CURRENT FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 10/002,796
PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18

PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31

PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27

PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21

PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25

PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16

PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/074092

PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25

PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08

PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10

PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18

PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17

PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24

PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28

PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20

PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23

PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15

PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20

PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17

PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07

PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19

PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29

PRIOR APPLICATION NUMBER: 09/114844
PRIOR FILING DATE: 1998-07-14

PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19

PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21

PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10

PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08

PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03

PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09

PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05

PRIOR APPLICATION NUMBER: 09/284663
PRIOR FILING DATE: 1999-04-15

PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14

PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14

STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKS, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF224
TELEPHONE: (301) 309-8509
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-275-805-2

Query Match 100.0%; Score 2407; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.9e-160;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 60
DB OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 60
QY 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 85
DB OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 120
DB TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 120
QY 86 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 145
DB TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 180
DB NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 180
QY 146 NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 205
DB NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 240
DB QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 240
QY 206 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 265
DB QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 265
QY 241 PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
DB PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
QY 266 PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
DB PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
QY 326 RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
DB RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
QY 361 GNEGREFYMRGTGPISATLVMTPTIKGPREIQLDLEMTVNTVINFRGSSVIRLRITYSQ 420
DB GNEGREFYMRGTGPISATLVMTPTIKGPREIQLDLEMTVNTVINFRGSSVIRLRITYSQ 420
QY 386 GNEGREFYMRGTGPISATLVMTPTIKGPREIQLDLEMTVNTVINFRGSSVIRLRITYSQ 445
DB GNEGREFYMRGTGPISATLVMTPTIKGPREIQLDLEMTVNTVINFRGSSVIRLRITYSQ 445
QY 421 YPF 423
DB 421 YPF 423
QY 446 YPF 448
DB 446 YPF 448

RESULT 4
US-09-836-561-1
Sequence 1, Application US/09836561
Patent No. US20020038006A1

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Corley, Neil C.
Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/836,561
FILING DATE: 16-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/212,168
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNNOT01
CLONE: 45517
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-836-561-1

Query Match 100.0%; Score 2407; DB 10; Length 448;
Best Local Similarity 100.0%; Pred. No. 8.9e-160;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 60
DB OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 60
QY 26 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 85
DB OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNGNGYLCIPRTNRYGPNPYS 85
QY 61 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 120
DB TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 120
QY 86 TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 145
DB TPYSGPYPAAPPLSAPNPTISRPLICRFGYOMESNOCVYDECATDSHCNPTQICI 145
QY 121 NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 180
DB NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 180
QY 146 NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 205
DB NTEGGYTCSTGTYWMLLEGQCLDIDECRYGCOQLCANVPGSYSCNPGFTLNEGRSC 205
QY 181 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 240
DB QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 240
QY 206 QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 265
DB QDVNECATENPCVOTCVNTYGSFICRDPGYELEEDGVHCSMDDECSFSEFLCOHCVNQ 265
QY 241 PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
DB PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 300
QY 266 PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
DB PGTYFCSGCPGYILLDDNRSCODINECEHRNHTCNLOQTCYNLQGGFKCIDPRCEEPYL 325
QY 301 RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
QY 326 RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
DB RISDNRCMCAENPGCRDQPTIILYRDMVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385

; INFORMATION FOR SEQ ID NO: 2;
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 448 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: No. US20020165151a1 Relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-10-041-016-2

Query Match 100.0%; Score 2407; DB 9; Length 448;
 Best Local Similarity 100.0%; Pred. No. 8.9e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIRPTNVPYRGPSNPYS 60
 |||||
 DB 26 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIRPTNVPYRGPSNPYS 85
 QY 61 TPYSGPYPAAPPLASAPNYPTISRPILCRFGYQMDENOCVVDDECATDSHCNPTQICI 120
 |||||
 DB 86 TPYSGPYPAAPPLASAPNYPTISRPILCRFGYQMDENOCVVDDECATDSHCNPTQICI 145
 QY 121 NTEGGYTCSTGTYWLEGGCLDIDECRYGTCQQLCANVPGSYSTCNGFTLNEDGRSC 180
 |||||
 DB 146 NTEGGYTCSTGTYWLEGGCLDIDECRYGTCQQLCANVPGSYSTCNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHCVNQ 240
 |||||
 DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHCVNQ 265
 QY 241 PGTYFCSPGTYILLDNRSQCQDINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 300
 |||||
 DB 266 PGTYFCSPGTYILLDNRSQCQDINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 325
 QY 301 RISDNRCMCAENPCRDQPTIILYRDMDVVSGRSVPADIFQMOATTRYPGAYYIFQIKS 360
 |||||
 DB 326 RISDNRCMCAENPCRDQPTIILYRDMDVVSGRSVPADIFQMOATTRYPGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPILKGPRIOLDLEMTVNTVINFRGSSVRLRLIYVSQ 420
 |||||
 DB 386 GNEGREFYMRQTGPISATLVMTRPILKGPRIOLDLEMTVNTVINFRGSSVRLRLIYVSQ 445
 QY 421 YPF 423
 |||||
 DB 446 YPF 448

RESULT 2
 US-09-083-002-2
 ; Sequence 2, Application US/09083002
 ; Patent No. US20010016650A1

; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Racine, Lisa A.
 ; APPLICANT: Lavallee, Edward R.
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Agostino, Michael
 ; APPLICANT: Lu, Zhijian
 ; APPLICANT: Honjo, Yasuko
 ; APPLICANT: Tashiro, Kei
 ; APPLICANT: Nakamura, Tomoyuki
 ; TITLE OF INVENTION: SECRETED PROTEINS
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 Cambridgepark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: U.S.A.
 ; ZIP: 02140

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/083.002
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sprunger, Suzanne A.
 ; REGISTRATION NUMBER: P-41,323
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8284
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 448 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-083-002-2

Query Match 100.0%; Score 2407; DB 10; Length 448;
 Best Local Similarity 100.0%; Pred. No. 8.9e-160;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIRPTNVPYRGPSNPYS 60
 |||||
 DB 26 QCTNGFPLDRSGGCLDIDECRTIPEACRGDMCMVONGGYLCIRPTNVPYRGPSNPYS 85
 QY 61 TPYSGPYPAAPPLASAPNYPTISRPILCRFGYQMDENOCVVDDECATDSHCNPTQICI 120
 |||||
 DB 86 TPYSGPYPAAPPLASAPNYPTISRPILCRFGYQMDENOCVVDDECATDSHCNPTQICI 145
 QY 121 NTEGGYTCSTGTYWLEGGCLDIDECRYGTCQQLCANVPGSYSTCNGFTLNEDGRSC 180
 |||||
 DB 146 NTEGGYTCSTGTYWLEGGCLDIDECRYGTCQQLCANVPGSYSTCNGFTLNEDGRSC 205
 QY 181 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHCVNQ 240
 |||||
 DB 206 QDVNECATENPCVOTCVNTYGSFICRCDPGYLEEDGYHCSMDSCSFSEFLCOHCVNQ 265
 QY 241 PGTYFCSPGTYILLDNRSQCQDINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 300
 |||||
 DB 266 PGTYFCSPGTYILLDNRSQCQDINECHRNHTCNLQOTCYNLGGFKCIDIRCEPYL 325
 QY 301 RISDNRCMCAENPCRDQPTIILYRDMDVVSGRSVPADIFQMOATTRYPGAYYIFQIKS 360
 |||||
 DB 326 RISDNRCMCAENPCRDQPTIILYRDMDVVSGRSVPADIFQMOATTRYPGAYYIFQIKS 385
 QY 361 GNEGREFYMRQTGPISATLVMTRPILKGPRIOLDLEMTVNTVINFRGSSVRLRLIYVSQ 420
 |||||
 DB 386 GNEGREFYMRQTGPISATLVMTRPILKGPRIOLDLEMTVNTVINFRGSSVRLRLIYVSQ 445
 QY 421 YPF 423
 |||||
 DB 446 YPF 448

RESULT 3
 US-09-275-805-2
 ; Sequence 2, Application US/09275805
 ; Patent No. US20010051358A1

; GENERAL INFORMATION:
 ; APPLICANT: OLSEN, HENRIK S.
 ; APPLICANT: LI, HAODONG
 ; TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HUMAN GENOME SCIENCES, INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:53:42 ; Search time 83.5316 Seconds

(without alignments)
582.319 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNFDLDRSGCCLDIDE.....INFRGSSVIRLRIVSQYF 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2407	100.0	448	9	US-10-041-016-2
2	2407	100.0	448	10	US-09-083-002-2
3	2407	100.0	448	10	US-09-275-805-2
4	2407	100.0	448	10	US-09-836-561-1
5	2401	99.8	448	9	US-10-066-500-15
6	2401	99.8	448	9	US-10-174-590-408
7	2401	99.8	448	9	US-10-176-758-408
8	2401	99.8	448	9	US-10-175-737-408
9	2401	99.8	448	9	US-10-173-706-408
10	2401	99.8	448	9	US-10-175-738-408
11	2401	99.8	448	9	US-10-175-752-408
12	2401	99.8	448	9	US-10-176-882-408
13	2401	99.8	448	9	US-10-176-913-408
14	2401	99.8	448	9	US-10-180-552-408
15	2401	99.8	448	9	US-10-180-557-408
16	2401	99.8	448	9	US-10-173-700-408
17	2401	99.8	448	9	US-10-174-572-408
18	2401	99.8	448	9	US-10-174-579-408
19	2401	99.8	448	9	US-10-174-579-408

20	2401	99.8	448	9	US-10-174-582-408
21	2401	99.8	448	9	US-10-174-588-408
22	2401	99.8	448	9	US-10-175-739-408
23	2401	99.8	448	9	US-10-175-740-408
24	2401	99.8	448	9	US-10-175-743-408
25	2401	99.8	448	9	US-10-176-488-408
26	2401	99.8	448	9	US-10-176-492-408
27	2401	99.8	448	9	US-10-176-747-408
28	2401	99.8	448	9	US-10-176-750-408
29	2401	99.8	448	9	US-10-176-985-408
30	2401	99.8	448	9	US-10-176-987-408
31	2401	99.8	448	9	US-10-176-991-408
32	2401	99.8	448	9	US-10-176-992-408
33	2401	99.8	448	9	US-10-176-993-408
34	2401	99.8	448	9	US-10-184-658-408
35	2401	99.8	448	9	US-10-002-796-15
36	2401	99.8	448	9	US-10-066-273-15
37	2401	99.8	448	9	US-10-066-494-15
38	2401	99.8	448	9	US-10-173-695-408
39	2401	99.8	448	9	US-10-173-697-408
40	2401	99.8	448	9	US-10-173-705-408
41	2401	99.8	448	9	US-10-174-576-408
42	2401	99.8	448	9	US-10-174-585-408
43	2401	99.8	448	9	US-10-174-586-408
44	2401	99.8	448	9	US-10-175-747-408
45	2401	99.8	448	9	US-10-176-481-408

ALIGNMENTS

RESULT 1
US-10-041-016-2
Sequence 2, Appl 1
Patent No. US2002016515A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
McCoy, John M.
Racle, Lisa A.
Lavallee, Edward R.
Merberg, David
Treacy, Maurice
Evans, Cheryl
Agostino, Michael
Lu, Zhijian
Honjo, Tasaku
TITLE OF INVENTION: SECRETED PROTEINS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/041,016
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/083,002
FILING DATE: 21-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: P-41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851

096SC3 PRELIMINARY; PRT: 2673 AA.
ID 096SC3;
AC 096SC3;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Fibulin-6 (Fragment).
GN FIBL-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Kostka G., Timpl R.;
RT "Partial sequence of fibulin-6 with a c-terminal region related to
RL domain II and III of the fibulin family."
DR EMBL; AJ306906; CAC37630.1; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000875; Cetrophin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF00008; EGF_5.
DR Pfam: PF00047; Ig_17.
DR Pfam: PF00090; tsp_1; 6.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_5.
DR PROSITE: PS00268; CECROPHIN; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_3.
DR PROSITE: PS01187; EGF_CA; UNKNOWN_8.
DR PROSITE: PS50092; TSP1; 6.
FT NON_TER 1
SQ SEQUENCE 2673 AA; 291017 MW; BEAEC30B8340E272 CRC64;
Query Match 22.5%; Score 542; DB 4; Length 2673;
Best Local Similarity 27.1%; Pred. No. 6.8e-46;
Matches 142; Conservative 65; Mismatches 195; Indels 122; Gaps 17;
QY 2 CTNFGDLDROSGGCLDIDECRTIPACRGDMVCVQNGGYLCIPRTNPYRGPNPST 61
DB 2170 CPKGLTIAADRTGCDIDECALGRHTCHAGQDCDNTISYRCVRCGSGFRRT-SDGLSC 2228
QY 62 PYSGPYRAAPPLAPNPPTTSRPLICRFQMDSESNOCVVDDECATDSHCNPTQICIN 121
DB 2229 QDINECOESSPCRHORCFNAISFHCCEPGYOL-KGRKCMVDNECR--QNVCRPDQCKN 2285
QY 122 TEGGYTC--SCTDGYWILE-GGCLDIDECRYGCO---QLCANYPGSYSCINPGFTLN 174
DB 2286 TRGYTKCIDLCPNKMTKAKENGTICDIDECCKDSTHOCRYNQCIENTRGSTRVCVCPRGYSQ 2345
QY 175 EDGRSCODVNEC-ATENPCVOTCVNTYGSFTICRCDPGYLEEDGVHCS----- 221
DB 2346 GVGRPCMDINECEQVPRKCAHGCSTPSPFCICPRGQHLIGDGKSCAGLERLPMNGTQY 2405
QY 222 ----- 221
DB 2406 SSYNLARFSPVNNYQPOOHYROYSHLYSSYSEYRNSRTSLSTRRTIRKTCPEGSEASH 2465
QY 222 ---DMDECSFSEFLCOHECVNPGTYFCSCPBGYILDDNRCODINECEHRNHTCNLQ 277
DB 2466 DTCVDIDECENTD-ACQHECKNTGFSYOCICPPGTQLTHNGKTCODIDECLEQNVHCGPN 2524
QY 278 QTCVNLGGGFKCIDPIRCEEPYLRIS-DNRGM--CPAENPGCRDQFTILYRDMVVSGR 334
DB 2525 RMCENMGSGYCID-TPCRPVYQRPVSGFCLKNCPNDLECALSPYALEYKLVSLPGSI 2583
QY 335 SVPADIFOMQATTR---YPCAYTI-----FOIKSGNEGREFYMRQGPISATLVM 381
DB 2584 ATNODLIRLVAYTODGVNHPRTTFLMVDDEQTVPFALRDEN-----LKGVVYT 2631

QY 382 TRPIKGR--EIQDLEMITVNTVINRGSSVIRLRIYVSOYP 423
DB 2632 TRPLREAEYRMVRASSANGTIEYQTFPIV--YLAVSAPY 2673

Search completed: July 3, 2003, 18:21:58
Job time: 32.0959 secs

```

OY 74 -LSAPNPTISRPICRFGYOMD-----ESNOCV-----DVDECATDSHOC 113
DB 298 CLNTPSPSFCIRLSCGTIAMDBSETRNNCFLLIINTFNCKYFEVEDVDECNLGSHDG 357
OY 114 NPTQICINTEGYTCS----- 129
DB 358 GPLYQCRNNGSYRCDAKCKGDELQNPMTGEYIDECVTGHNCGAGECVMTPGSRQOQ 417
OY 130 -----CIDGYWL--LEGGCLDIDECRYGCOOL-CANPGASISCTCNPGFTLNE----- 175
DB 418 KGNLCAHGEVNGATGECEDYNNEQDQGVCSMECINLPGYKCKGCGPEYERNDARKCED 477
OY 176 ----- 197
DB 478 VDECIFAGHYCDLSAECINTIGSFECCKRGPOLASDGRCEDEYNECTTGACGEOKCV 537
OY 198 NTYGSFICRCDPGYELEEDGVHCSMDCECF-----SEFLCOHECVNPGYFCSCPGRYI 253
DB 538 NIPGSYOCICDRGFALGPDGTCKEDIDECISWAGSGNDLCKMGCCINTKSGYLQCPGPYK 597
OY 254 LDDNSCODINCEHNRHTCNLOQTCVNLQGGFKCIDPRICEPYLAIISDR----- 306
DB 598 IOPDGRTCVDVDECA-MGECAGSDKVCVNTLGSFKC-HSIDCPYNIHDSLNRKNCNQP 655
OY 307 --CMCPAENPGCHDOFTILYRDMDVYSGRSVP-----ADI---FOM 343
DB 656 SAGCLPEE--CKVPLFLTYQFISL--ARAVPISSHRPATLTKRVASPNADTEVNEEL 710
OY 344 QATRRYPAYIYIQISGNEGREFYMQTGP--ISATLVMTRPIKPREIOLDL 395
DB 711 QKTTIVGAPNVLPATRA-----FLLOKGEKRRSAVVTLRDSDGPOTVKQL 759

RESULT 14
OYVS89 PRELIMINARY; PRT; 1394 AA.
AC 09VS89;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE CG7526 protein (Fragment).
GN CG7526.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
(1)
SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agapayni A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.M., Nelson K.A., Nixon K., Nussken D.R., Pacle V.J.M.,
RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhan M., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003558; AAF50538.1; -.
DR HSSP: P00736; IAPQ.
DR InterPro: FBgn0035798; CG7526.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003410; Hyalin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001491; Thrombomodulin.
DR Pfam: PF00008; EGF_11.
DR Pfam: PF00084; sushi_2.
DR PRINTS: PR00907; THROMBOMODULN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00179; EGF_CA_9.
DR SMART: SM00001; EGF_Like_5.
DR PROSITE: PS00010; ASX_HYDROXYL_8.
DR PROSITE: PS01187; EGF_CA_10.
DR PROSITE: PS01187; EGF_CA_9.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW NON_TER 1394 1394
FT SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;
SQ

```

```

Query Match 22.7%; Score 547.5; DB 5; Length 1394;
Best Local Similarity 32.7%; Pred. No. 8.6e-47;
Matches 119; Conservative 40; Mismatches 128; Indels 77; Gaps 13;

```

```

OY 2 CTNGFD--LDROGOCIDIDECRTIPEACRGDMQVQNGSYLCIPRTNPPYRGPSNP 58
DB 519 CAGGYELKIDGIRGCFDIDECGQRTHGCGSDQMLCENLNGSYTCL--CPRGYALGLDH 576
OY 59 YSTPYSGPYPAAPPLSAPNPTISRPICRFGYOMDSNOCVDYDECATDSHOCNPTOI 118
DB 577 IVTSLNSSFYTDSTSSSETPS-----AHCTCIDIDECSLANGNC--SHR 616
OY 119 CINTREGYTCSCDTGMYLLEGO--CLDIDEC--RYGYCOQLCANVPGSYSCCNPGFTLN 174
DB 617 CONEPGFGQACPLGYALISEMRKCODIDECIDLSNGQCSQLCLNPPGFCACETGFELT 676
OY 175 EDGRSCDVCACATE-NPCVQTCVNTYGSFTCRCDPGYELEEDGVHCSMDCECS-FSEPL 232
DB 677 PDGFGCADIDECSDGYNCNSDICNLGLTHACACRGYELAKDKLISCLDVDECAGLSLGG 736
OY 233 COHECVNQPGYFCSCPGRYILLDNNSCQ----- 262
DB 737 CSHECINKAATFECCGCPGLYLLNDGRCSALVWGPRTORSADGCAPIECNPGYTLGS 796
OY 263 -----DINEGHRHTCNLOOTCVNLGGFCRCIDPRICEPYLRIISDR-CK-----CPAE 312
DB 797 DDXCVYDCEQKONGGCS--HRCNSTESSFC--SCPPGYELSDQTKCODIDECDD 850
OY 313 NPGC 316
DB 851 KTSC 854

```

RESULT 15

```

Db      458 DSLNKNQJADGYSCKIVCTEDTECGNTREVLYFRAVPSPKTTISPLEVSRIVTHMG 517
QY      341 --FOQATRIYGAAYIIFOIKSGNEGREYMGOTGPISATLVTRPIKGPRTOLDLEMI 398
Db      518 VPFSDVMDLVGGYGRHFRVJDERNIG-----IYLVKVISGP-----IVERTI 559
QY      399 TVNTVINFSGSSVIR-----LRIVYSOYPF 423
Db      560 KVNHTKSRGTGVIILFNAEIIITSVSKYF 589

RESULT 12
ID      090H16      PRELIMINARY;      PRT;      554 AA.
AC      090H16;
DT      01-MAY-2000 (Tremblrel. 13, Created)
DT      01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      D162H14.1 (Fibulin 1) (Fragment).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Lloyd D.;
RL      Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Z98047; CAB6295.1; -.
DR      HSP; P35555; IEMN.
DR      InterPro; IPR000020; Anaphylatoxin.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      Pfam; PF01821; ANATO; 3.
DR      Pfam; PF00008; EGF; 5.
DR      SMART; SM00104; ANATO; 3.
DR      SMART; SM00179; EGF_CA; 4.
DR      SMART; SM00001; EGF_Like; 5.
DR      PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR      PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR      PROSITE; PS00010; ASX_HYDROXYL; 4.
DR      PROSITE; PS01186; EGF_2; 3.
KW      EGF-like domain; glycoprotein; Hydroxylation.
FT      NON_TER
SQ      SEQUENCE      554 AA; 59767 MW; FE285184599A2982 CRC64;

Query Match      24.2%; Score 581.5; DB 4; Length 554;
Best Local Similarity 36.7%; Pred. NO. 9.7e-51;
Matches 114; Conservative 46; Mismatches 100; Indels 51; Gaps 12;

QY      1 OCTNGFDLROSGQCIDIDECPRTIPACRGDMCVNONGYLIPRTNPNYRGYSPY 60
Db      266 QCKSGFIQD-ALGNCIGINECLISAPCRIGHCINTEGSYTC----- 307
QY      61 TPYSGPYRAAPLSPNPTTISRPLICRGYQM-DESNQCVDYDECATDSHOCNPTQIC 119
Db      308 -----QKNVFN-----CGRGYHLNEGTRCYGVDECAPAPAEPCGKGRHC 346
QY      120 INTEGGYCSCTDGYWL--LEGQCIDIDBC-RY--GYCQOLCANVGSISCTNPGFTLN 174
Db      347 VNSPFSFCECKTGYFIDGISMKCVGNECQRPGRLCGHKCENTLGSYLCSGSGFRLS 406
QY      175 EDGRSCQDVNECATENPCVQTCVNTYGSFICRCDDPEYLEE-DGVHCSMDDECSF--SEF 231
Db      407 VDGRCCEGIEGSS-SPSQSQCANYGSGYQCGRGYQSLDVGTYCEGIDECALPTGGH 465
QY      232 LCGHCVNQGTYFPCSP-PTGILLDDNNSCODINECEHRNHTCNLQQTCTYNLOGGFKCI 290
Db      466 ICSYRCLNIPSSFGSCSSPSGRILAPNGRCOIGIDECVYGINCSINTECFNIGGFRCL 525
QY      291 DPICEPEYLR 301

```

```

Db      526 -AFECPEMYR 535

RESULT 13
ID      018026      PRELIMINARY;      PRT;      798 AA.
AC      018026; Q20903;
DT      01-JAN-1998 (Tremblrel. 05, Created)
DT      01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT      01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE      F56H11.1 protein.
GN      F56H11.1.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitoidea;
OC      Rhabdilitae; Pelodierinae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Wilkinson J.;
RL      Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RX      MEDLINE=94150718; PubMed=7906398;
RA      Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA      Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA      Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA      Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA      Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA      Lightning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rikken L., Roopra A., Saunders D., Showkhen R.,
RA      Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA      Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
RA      Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
RA      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT      elegans."
RL      Nature 368:32-38(1994).
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Lloyd C.;
RL      Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Z68749; CAA92962.1; -.
DR      EMBL; Z68219; CAA92962.1; JOINED.
DR      EMBL; Z68219; CAA92483.1; JOINED.
DR      HSP; Z68749; CAA92483.1; JOINED.
DR      HSP; P16109; IESB.
DR      InterPro; IPR000020; Anaphylatoxin.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR001881; EGF_CA.
DR      Pfam; PF01821; ANATO; 2.
DR      Pfam; PF00008; EGF; 5.
DR      SMART; SM00104; ANATO; 2.
DR      SMART; SM00179; EGF_CA; 4.
DR      SMART; SM00001; EGF_Like; 6.
DR      PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR      PROSITE; PS00010; ASX_HYDROXYL; 4.
DR      PROSITE; PS01186; EGF_2; 5.
DR      PROSITE; PS01187; EGF_CA; 8.
KW      Calcium-binding; EGF-like domain; glycoprotein; Hydroxylation; Repeat.
SQ      SEQUENCE      798 AA; 87205 MW; 3BF1EE9ED54DBB9 CRC64;

Query Match      23.8%; Score 574; DB 5; Length 798;
Best Local Similarity 26.8%; Pred. NO. 8.7e-50;
Matches 159; Conservative 49; Mismatches 172; Indels 214; Gaps 22;

QY      2 CTNGFDLROSGQCIDIDECPRTIPACRGDMCVNONGYIC-----IP-----RTNPNY 51
Db      180 CRSGFDLAPGMACVDRNECLTROSPTQSEDCVNTIGGYICQRIISRLVPHRRRAVIG 239
QY      52 RGP---YSNPYSTPYSGPYPAAP----- 73
Db      240 NAFRRMRDDPYSR-AGEYREASQANTEFCGPMGWLRFQHGCVYDIDECATLMDCLSOR 297

```


RESULT 10
 ID 042182 PRELIMINARY; PRT: 681 AA.
 AC 042182;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Fibulin-1 D.
 GN FBLN1.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 ON NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang H.-Y., Lardelli M., Ekblom P.;
 RT "Sequence of zebrafish fibulin-1 and its expression in developing
 heart and other embryonic organs.";
 RL Dev. Genes Evol. 0:0-0(1997).
 DR EMBL: AF013751; AAB80944.1; .
 DR HSSP: P35555; 1EMN.
 DR ZFIN: ZDB-GENE-990415-73; fbln1.
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR Pfam: PF01821; ANATO. 1.
 DR Pfam: PF00008; EGF. 4.
 DR SMART: SM00104; ANATO. 2.
 DR SMART: SM00179; EGF_CA. 5.
 DR SMART: SM00001; EGF_Like. 4.
 DR PROSITE: PS01177; ANAPHYLATOXIN.1; UNKNOWN_1.
 DR PROSITE: PS01178; ANAPHYLATOXIN.2; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL. 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA. 6.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW SEQUENCE 681 AA; 74459 MW; 175C966305A46699 CRC64;
 SQ

Query Match 26.28; Score 629.5; DB 13; Length 681;
 Best Local Similarity 32.88; Pred. No. 1.5e-55;
 Matches 151; Conservative 64; Mismatches 160; Indels 85; Gaps 21;

1 QCTNGFDLDROSGOCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNPNVYRGPSNPYS 60
 270 QCAAGFIQD-ALGSCIDINECVSYTALSRS-QMCFNTYGSFTICQHS----- 314
 61 TPYSGPYPAAPLPSADPNYPTISRPILICRFGYOMD-ESNOCVVDDECATDSHCNPOTIC 119
 315 -----VTCGRGYHINAEGTRKCVIDDECAGPDNSCD-GHGC 348
 120 INTEGGYTCCTDGYWL-LEGCLDIDECRYG-----COOLCANVPSSYSTCNPFTL 173
 349 INLVGSTRCCRGTFENISRSCEIDECR-NYPGRCAHKCENILLISYKSCYAGFKL 407
 174 NEDGRSCDDNECATENPCYQTCVNTYGSFTICRCDPELEE-DGVHCSMDDECSF--SE 230
 408 ADGRNCDDVNEES-SPCSGCANVGSYSTYCRGQLSDADGITEDEIDECALPIGCG 466
 231 FLCQHCNCPGTYFCSCP-PGYILLDNRSQDINECEHRNHTCNLOQTCYNLGGFRC 289
 467 HICSYCHNTPGSGFHTCTCPASGYTLAANGRSQDIDELTGTSCSESESCFNIGGFR 526
 290 IDPITCEEPYLRISDN-----RCM--CPAENPCGRDQPTILYLRDMVVS----- 332
 327 LS-PDCPANTRRGDTRPRVDRADITRCVASCQNDNSCVLP--TISHSHTAISLPFR 583
 333 GRSVPADIFOMQATT-----RYPGAYYIFQIKSGNEGEFFYMRQ---GPISATLWMTRP 384
 584 EFKRPEIYFLRBPPTPLPHMDSPELYVDLEGNINQSPILIKRLDGMIVGVKQYRP 643
 385 IKGREIQLDLEM-ITVNTVINRGSSVIRLRIYVSQTF 423

Db 644 LVGPFRTVLKLANMYTVNGVYSHR--NIINRIYSEFWF 681
 RESULT 11
 ID 097251 PRELIMINARY; PRT: 589 AA.
 AC 097251;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Fibulin-ID (Fragment).
 GN FBLN1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodietidae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CB1489 HIM-8;
 RX MEDLINE=99120531; PubMed=9923656;
 RA Barth J.L., Argaves K.M., Roark E.F., Little C.D., Argaves W.S.;
 RT "Identification of chicken and C. elegans fibulin-1 homologs and
 characterization of the C. elegans fibulin-1 gene.";
 RL Matrix Biol. 17:635-646(1998).
 DR EMBL: AF070477; AAC24035.1; .
 DR HSSP: PL6109; IFSB.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-Like.
 DR InterPro: IPR001881; EGF_CA.
 DR Pfam: PF00008; EGF. 6.
 DR SMART: SM00179; EGF_CA. 5.
 DR SMART: SM00001; EGF_Like. 4.
 DR PROSITE: PS00010; ASX_HYDROXYL. 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_CA. 7.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
 KW NON-TER 1
 FT SEQUENCE 589 AA; 63984 MW; 8EA3E8FC0B97BE6 CRC64;
 SQ

Query Match 24.68; Score 592.5; DB 5; Length 589;
 Best Local Similarity 27.48; Pred. No. 7.8e-52;
 Matches 156; Conservative 52; Mismatches 155; Indels 207; Gaps 20;

2 CTNGFDLDROSGOCLDIDECRTIPEACRGDMCVNONGGYLCIPRTNPNVYRGPSNPYS 60
 79 CRSGFDLAPGMAVCVDHIDECATLMDCLBESQRCINTGFSFCT----- 122
 61 TPYSGPYPAAPLPSADPNYPTISRPILICRFGYOMD-ESNOCVVDDECATDSHCNPOTIC 119
 123 -----RTLSCGTGYAMDSFERCRDVEDCNLGSDCPLVQC 159
 120 INTEGGYTC-----SCTDGYWLLEGCLDIDECRYG----- 151
 160 RNTQGSYRCDAKKCGDELQNPMTGECTSYICPNGYPRKNGCNDIDECVTHNCGAGE 219
 152 -----COO--LCA-----NVPGSYSCNPG 170
 220 CVNTPGSRQDQKGNLCAHGEVNGATGFCEDVNECCQGVGSGSMECTINLPRTYKCKCGG 279
 171 FTLINE-----DGRSCDDVNEC 186
 280 YEFNDARKCEDVEDECIKFAGHYVCDLSAECINTIGSFCKCKPGFQLASDGRRCEDVNEC 339
 187 ATE-NPCVQCVNTYGSFTICRCDPELEEBCVNCSDMDECSF-----SEFLCQHCNVP 241
 340 TTGIAACEQKCVNIPGSGYICIDGFGALPGGTGCEIDECSTIMAGSNDLGMGCINTK 399
 242 GTFYSCGPYIILLDNRSQDINECEHRNHTCNLOQTCYNLGGFCKIDPITRCEEPYLR 301
 400 GSTLCQCPPTKIDPDRCTCYDVDECA-MGCASADSKVCVNTIGSFCK-HSIDCTNTYIH 457
 302 ISDNR-----C-MCPAENPCG-RDQPTILYLRDMVVSGRSV--PADI----- 340

```

Db 554 S-LECPENYRKSGDVTLEKTDITRCIKSGRPNDVNCVDPVHTISHTVLSLPTFERETR 612
QY 337 PADIFQMOA-TTRYPG--AYITQISGNGREF-----YKQTGPISATLVMPRIKGP 388
Db 613 PEELIFLRAITTPYANQADIIDITEGNRESFDIKRYM--DGMTVGVVQVRRPIVGP 670
QY 389 REIQLDLEM-ITVNTVINFRSSVIRLIYVSOYPF 423
Db 671 FHALIKLENNYMGVYSHR--NIVNHTVSEIWP 704

RESULT 8
QY3V7 PRELIMINARY; PRT; 576 AA.
ID 09Y3V7;
AC 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical 63.3 kDa protein (Fragment).
NM DKFZP586A1519.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL050095; CAB43267.1; -.
DR HSSP: P00736; IAPQ.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_6.
DR SMART: SM00179; EGF_CA; 8.
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA; 9.
DR Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation;
KW Hypothetical protein; Repeat.
FT NON_TER
SQ SEQUENCE 576 AA; 63274 MW; FF2F0E9B185D8AC3 CRC64;

Query Match 29.2%; Score 703.5; DB 4; Length 576;
Best Local Similarity 34.5%; Pred. No. 3.5e-63;
Matches 142; Conservative 58; Mismatches 152; Indels 59; Gaps 11;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNPPYRGPSNPYS 60
Db 187 RCMDSF-LQDPGNCVDINECTSLSPCRPFSCTINTVSGYT----- 228
QY 61 TPYSGPYAAAPPLSAPNPTISRLPICRFYQ-MDESNQCVNDECATDSHOCNPYQIC 119
Db 229 -----QRNPLICARGYHASDGGKCVNDEGTEVHHGGEQVC 267
QY 120 INTEGTCSCSTDGWM--LLEGQCLDIDECRG--YCOOLCANVGSYSCCTNPGFTLN 174
Db 268 HNLPSYSCDCXAGQFQDAFGRGICIDVNCMASPGRLCOHTCENTLGSYSCASGFLTA 327
QY 175 EDGRSCQDVNCEATENPCVQTCVNTYGSFICRCDDPEYELEEDGVHSCDMECS-FSEFLC 233
Db 328 ADGRKCEVNECEAOR-CSQECANITYGSYQCYCRGQYQLAEDGHTCTDIDECAGAGAILC 386
QY 234 QHECVNQGTYFCSCP-PGYILLDNRSQDINECHRNHTCNMQTCYMLQGGFKCIDP 292
Db 387 TFRCLNVGSGYQACPEQGYTWTANGRSCKVDDECALGTHNCEASEATCHNIGGSFRL-R 445
QY 293 IRCEEPYIRISDNRCMCPAENPCRD-----QPTILYRDMVDVSGRSVPADIFQMQAT 346
Db 446 FCEPNNYVQVSKTKC-----ERTTCHDFLECONSFAITHYQINFORGLVPAHIFRIGPA 501

```

```

QY 347 TRPGAYYIFQIKSGNEGRFYKRGTPISATLVMPRIKGPRIQLDLEM 397
Db 502 PAFGTGTLALINTIKGNEGEYFGRRLNAYGVVYLQRAVLEPRDFALDVEW 552

RESULT 9
QYH05 PRELIMINARY; PRT; 495 AA.
ID 09H05;
AC 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 54.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qiu W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF217999; AAG17241.1; -.
DR HSSP: P35555; IENM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF_6.
DR SMART: SM00181; EGF_9.
DR SMART: SM00179; EGF_CA; 9.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 8.
KW EGF-like domain; Glycoprotein; Hydroxylation; Hypothetical protein.
SQ SEQUENCE 495 AA; 54340 MW; C40434B6C82E3370 CRC64;

Query Match 28.0%; Score 674.5; DB 4; Length 495;
Best Local Similarity 36.2%; Pred. No. 2.7e-60;
Matches 137; Conservative 58; Mismatches 126; Indels 57; Gaps 15;

QY 1 OCTNGFDLDRSGGCLDIDECRTIPACRGDMVCVNONGYLCIPRTNPPYRGPSNPYS 60
Db 163 QCKSGFIOD-ALGNICIDINECTSLSAPCEIGHTCINTWEGSYTC----- 204
QY 61 TPYSGPYAAAPPLSAPNPTISRLPICRFYQ-MDESNQCVNDECATDSHOCNPYQIC 119
Db 205 -----QKNVFN-----CGRGYHLINEGTRCVNDECAPPAEECGKGRHC 243
QY 120 INTEGTCSCSTDGYWL--LLEGQCLDIDEC-RY--GYCOOLCANVGSYSCCTNPGFTLN 174
Db 244 VNSPSFCECKTGYFEDISRMCVNDEGTCQRPGLCKHCKGNTIGSYLCSGSGVGRRLS 303
QY 175 EDGRSCQDVNCEATENPCVQTCVNTYGSFICRCDDPEYELEF-DGVHSCDMECSF--SEF 231
Db 304 VDRGSCEDINECS--SPSOECANVGSYQCYCRRGYQSLVDGVNCEIDECALPTGGH 362
QY 232 LCOHCEVNOGTYFCSCP-PGYILLDNRSQDINECHRNHTCNMQTCYMLQGGFKCI 290
Db 363 ICSTRCLNIPGSFQSCPSGSRRLAPNGNCDIDECVTGHNCSINTECFNIGGFRCL 422
QY 291 DPIRCEEPYIRISDNRC-MCPA-ENPGCRDQPTILYRDMVDVSGRSVPADIFQMQAT 348
Db 423 -AFDEPENYRRAAARICERLPCHENRECKLPLRTIYHLSPTNQAARAVYFRMGPSA 481
QY 349 YPGAYYIFQIKSGNEGR 366
Db 482 VPGD-----SMQAGHHRQ 495

```



```

0Y 61 TPYSPYPAAPPLSPNPTISRPLICAFEGYOMDESNOCVADDECATOSHOCNPJOICI 120
Db 94 ---EGP-PPYVPAQHFN-----PCPPGTGEPDDQSCVDVDECAQALHDBRRPSODCH 141
0Y 121 NTEGGYTCSDTDGYMLLEGQCLDIDECRAYCCOQLCANPGYSCTCNFGFTLNEBGRSC 180
Db 142 NLSGSGYOGCTCPDGYKRIIGPECYVDIDECRRYCOHQHRCVNLPGSRCCGCEFGQGPNNRSC 201
0Y 181 ODVNECCATENPCVOTCVNTGSEFICRCDPGYELEEDGYHCSMDDECSFSEFTLCQHECVNO 240
Db 202 VDVNECDMDGAPCEORCFNSGYTFELCRCHOGYELHRDGFSCSDIDECSSSYLQCYRCVNE 261
0Y 241 PGTYECSCEPGYIILLDDNRSODINCEHRNHTCNLDQOYCNLGGCFKCIDPIRCEBPLY 300
Db 262 PERECSHCDCQIQL-ATRLCODIDECESGAHOCSEAOYTCVNHGGYRCVDNIRCVBPTY 320
0Y 301 RISDNRCMCPAENPGCRDQPFITILYKRDMDVVGSRSPADIFQWQATTBPYGAYYFOIKS 360
Db 321 OYSENRCLCPASNPLCRBGPSSIVHRHYMTITSEBMPADAVFOQAATSVYFGAIAINAQOIRA 380
0Y 361 GNEGHEFTYRQOTGPISATLVMTRPILKGPRIQLDEMTVTNVTYNINFGRSVILRLIYVSO 420
Db 381 GNSQGDFTYRQINNTVSAMLVLARPYTGPPEVYIDLEMTMNSLMSYRASVLRLLTYEVGA 440
0Y 421 YPF 423
Db 441 YTF 443

```

RESULT 4	ID	0922K8	PRELIMINARY;	PRT;	685 AA.
AC	0922K8	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Similar to fibulin 1.				
GN	FBIN1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_Taxid=10090;				
RM	[1]				
RP	SEQUENCE FROM N.A.				
RA	Strausberg R.;				
RL	Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC007140; AA07140.1; -				
DR	MGD; MGI:95487; Fbln1.				
DR	InterPro; IPR000020; Anaphylatoxin.				
DR	InterPro; IPR000152; Asx_hydroxyl.				
DR	InterPro; IPR000561; EGF-like.				
DR	InterPro; IPR001881; EGF_Ca.				
DR	Pfam; PF01821; ANATO; 3.				
DR	Pfam; PF00008; EGF; 6.				
DR	PROSITE; PS01177; ANAPHYLATOXIN_1; UNKNOWN_3.				
DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 3.				
DR	PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_4.				
DR	PROSITE; PS01186; EGF_2; UNKNOWN_3.				
DR	PROSITE; PS01187; EGF_CA; UNKNOWN_8.				
SO	SEQUENCE 685 AA; 75283 MW; EF0D77D7F66B73B8 CRC64;				
Query Match	30.8%; Score 741.5; DB 11; Length 685;				
Best Local Similarity	36.5%; Pred No. 5.7e-67;				
Matches 158; Conservative 64; Mismatches 152; Indels 59; Gaps 16					
QY	1 QCTNGFDLDROSGCLDIDECRTIEACRGDMKCYNONGYLICPTPTNPFYRGPSNPYS 60				
Db	295 QCKSGFIQD-ALGNCIDINECLISAPCVGQTCINTBSYTC----- 336				
OY	61 TPYSGFYAAAPPLSPAPNPTITSRLICFRGYQM-DESNQCVYDVECDATDSHOCAPTQLC 119				
Db	337 -----QKNVPN-----GGRGHYLINEGRVCDYDECCSPAEPCGKGHHG 375				

```

QY 120 INTEGGYTCSDJGD/YWL--LEGGCIDLDEC-RY--GYCQQLANVAGSYSCOTNCPPTLL 174
Db LNSPSPFRCCEAGYAFDGIISFTVCIDINCCORPGRULCCHKCKENTGSPHSCSAGFRLS 435
QY 175 EDGRSCQOVNCECATENPCVQTCVNTYGSFIFCRADPGEYLEE--DGVHCSMDDECSE--SEP 231
Db 436 VDRGRCEDEVNEC-LNSPSCQBEACANYAGSYQTCYCRSTQJLSDVDGVTCEIDICALPTTGGH 494
QY 232 LQNHBCVNOQPTGYFFCSCP-PGYILLDDNMSCODINECEHRNHTCNLQOTCYNLQGGFKCI 290
Db 495 ICSYRCINPMSFGQCSPPSSGYRLAPNGRNQDIDECVYGIHNCISINTECFNIQGSFRCL 554
QY 291 DPICEEPTYLRLSDNRC-MCPA--ENPGCKDQPTILYRMDVYSGHVSVPADIFQQAITR 348
Db 555 S-FECPENYRRSADRCERLPCHENOECRPLRLIRTYHGLSEPTNIQVAVAFVRMGPSSA 613
QY 349 YPGAVYINQKSGNGREPYMRQTOGTSILMTMPRIKAPREQLDLEN--ITVNVVINF 406
Db 614 VPGDSMQLATITAGNEEGFTTRKRVSHSHSEVALTKPIPEPRDLLLTVMKMDLYRHGTVSSP 673
QY 407 RGSYSIRLRIYVS 419
Db 674 ----VAKLFIFVS 682

```

RESULT 5	ID	Q8TBH8	PRELIMINARY:	PRT:	683 AA.
AC	Q8TBH8:				
DT	01-JUN-2002 (TREMBLrel, 21, Created)				
DT	01-JUN-2002 (TREMBLrel, 21, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel, 21, Last annotation update)				
DE	Fibulin 1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
NC	NCBI Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RA	Strausberg R.;				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: BC022497; AAH22497.1; --				
SQ	SEQUENCE 683 AA: 74423 MW; 2665A3961B6403B4 CRC64;				
Query Match					
Best Local Similarity	30.6%; Score 735.5; DB 4; Length 683;				
Matches	155; Conservative 66; Mismatches 152; Indels 61; Gaps 16;				
QY	1 QCTNGFDLDROSGCCLDIDECRTTPEACRGDMCVNONGYLCTIPRTNPNYRGPSNPYS 60				
DB	293 QCKSGFID-D-ALGNCIDINECLISAPBCPIGHTCINTGSGYTC----- 334				
QY	61 TPYSGPYPAAPRPSAPNPTISRLPICRFGYOM-DESNOCVYDDECATDSHOCNPQIC 119				
DB	335 -----QKNVPR-----CGRGYHNEBETRCYVDDECAPRABPCGKGRHC 373				
QY	120 INTEGGYCTSGTDYWT--LEGGCLDIDEC-RY-GYCOQLCANYPGSGYSCTCNPFPTLN 174				
DB	374 VNSGSGFCECKTGTYFDGDISRMCDVNECQRYRPERLGGHNCENTLGSYLCSSVGFRLS 433				
QY	175 EDGSSCODVNCATLENPCVQTCVNTYGSFGTCRCDPGVYLEE-DGYHGSMDMECSR--SEF 231				
DB	434 VDGSSCEDINECSS-SPECSQECANVYGSQCYCRGQYQLSPVDGYTCIDDECAIPTGGH 492				
QY	232 LCOHEACVNOPTGYFCSCP-PGYIILDDNRSCODINECEHNRHTCNLAQTCYNLGGFKIC 290				
DB	493 ICSYRCINIPESFQCSQSSSGYRLAPNSNODIDECYTGHNCSINTCFNICGGFRLC 552				
QY	291 DPICEEPEYLISDNRC-MCPA-ENPGCRDOPFTILYRDMDVVSGRSVPADIFOMQATTR 348				
DB	553 -AFCEPEYRRSAAATRCERLPCHEHNRCSKLRPLRTYYHLSEFPNIIQAPAVAVFTRMGPSA 611				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 3, 2003, 15:34:32 ; Search time 30.5959 Seconds
(without alignments)
2848.683 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFDLDROSGCCLDDE.....INFRGSSVIRLRIVSYQYF 423

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

otal number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organella:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertelbrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriaph:*
- 17: sp.archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1276.5	53.0	443	4	Q96FT5
2	1273.5	52.9	443	11	Q9JMK6
3	1253.5	52.1	443	4	Q9H3D5
4	741.5	30.8	683	11	Q922K8
5	735.5	30.6	683	4	Q8TBH8
6	731.5	30.4	1174	11	Q99K58
7	713	29.6	704	13	Q73774
8	703.5	29.2	576	4	Q9Y3V7
9	674.5	28.0	495	4	Q9HBQ5
10	629.5	26.2	681	13	Q42182
11	592.5	24.6	589	5	Q9RTS1
12	581.5	24.2	554	4	Q9UHI6
13	574	23.8	798	5	Q18026
14	547.5	22.7	1394	5	Q9VSB9
15	542	22.5	2673	4	Q96SC3
16	539	22.4	1394	4	Q8TD95

17	531	22.1	5636	4	Q96RW7	Q96RW7 homo sapien
18	518.5	21.5	2872	11	Q9WUH8	Q9WUH8 rattus norv
19	517.5	21.5	741	4	Q96K89	Q96K89 homo sapien
20	512.5	21.3	3857	11	Q88840	Q88840 mus musculu
21	512	21.3	1713	11	Q88349	Q88349 mus musculu
22	511.5	21.3	2809	4	Q96UP8	Q96UP8 homo sapien
23	508	21.1	188	11	Q8RIU8	Q8RIU8 mus musculu
24	507.5	21.1	708	13	P87363	P87363 gallus gall
25	506.5	21.0	2906	11	Q9WUH9	Q9WUH9 rattus norv
26	491.5	20.4	937	5	Q9BR11	Q9BR11 clone intes
27	484.5	20.1	1963	6	Q28019	Q28019 bos taurus
28	483.5	20.1	447	5	Q9VAB8	Q9VAB8 dtrosophila
29	482.5	20.0	1764	11	Q35806	Q35806 rattus norv
30	480.5	20.0	517	4	Q9NP01	Q9NP01 homo sapien
31	474	19.7	746	4	Q96HB9	Q96HB9 homo sapien
32	474	19.7	1256	4	Q9NS15	Q9NS15 homo sapien
33	474	19.7	1382	4	Q9H7K2	Q9H7K2 homo sapien
34	471.5	19.6	1095	11	Q60784	Q60784 mus musculu
35	469.5	19.5	1511	4	Q75412	Q75412 homo sapien
36	469.5	19.5	1664	5	Q9TVQ2	Q9TVQ2 caenorhabd1
37	469.5	19.5	1821	4	Q14767	Q14767 homo sapien
38	468.5	19.5	669	4	Q75441	Q75441 homo sapien
39	468.5	19.5	1062	11	Q60789	Q60789 mus musculu
40	467	19.4	1289	5	Q8SSS3	Q8SSS3 dictyosteli
41	466	19.4	1587	4	Q00508	Q00508 homo sapien
42	465.5	19.3	2189	5	Q9BI05	Q9BI05 elmeria ten
43	462.5	19.2	733	5	Q25678	Q25678 podocoryne
44	460.5	19.1	961	11	Q9BQC6	Q9BQC6 mus musculu
45	455	18.9	937	4	Q96FT5	Q96FT5 homo sapien

ALIGNMENTS

RESULT 1

ID Q96FT5 PRELIMINARY; PRT; 443 AA.

AC Q96FT5;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Mutant p53 binding protein 1 (MBP1).

GN MBP1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Tanaka S.;

RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.

RN (2)

RP SEQUENCE FROM N.A.

RA Tanaka S., Sugimachi K., Sugimachi K.;

RT "Human mutant p53 binding protein (MBP1).";

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AB030655; BAA2880.1; .

DR InterPro: IPR000152; Asx_hydroxyl.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR001881; EGF_Ca.

DR Pfam: PF00008; EGF_4.

DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_4.

DR PROSITE: PS01186; EGF_2; UNKNOWN_4.

DR PROSITE: PS01187; EGF_CA; UNKNOWN_6.

SQ SEQUENCE 443 AA; 49421 MW; 9CE175FAF388A56D CRC64;

Query Match 53.0%; Score 1276.5; DB 4; Length 443;

Best Local Similarity 52.0%; Pred. No. 8e-122;

Matches 220; Conservative 67; Mismatches 119; Indels 17; Gaps 4;

OY 1 OCTNGFDLDROSGCCLDDECRITPEACRGDMCYNONGYLCIRTPNVPYSGPSNPTS 60

DB 38 ECTDGYEMDPDSQHCROVDNECLITPEACRGEMKICINHYGYLCILPRAAIVNDLHG---- 93

FT DISULFID 640 654 BY SIMILARITY.
FT DISULFID 657 669 BY SIMILARITY.
FT DISULFID 675 686 BY SIMILARITY.
FT DISULFID 681 695 BY SIMILARITY.
FT DISULFID 698 709 BY SIMILARITY.
FT DISULFID 715 726 BY SIMILARITY.
FT DISULFID 721 735 BY SIMILARITY.
FT DISULFID 737 750 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 762 776 BY SIMILARITY.
FT DISULFID 778 791 BY SIMILARITY.
FT DISULFID 797 808 BY SIMILARITY.
FT DISULFID 803 817 BY SIMILARITY.
FT DISULFID 819 832 BY SIMILARITY.
FT DISULFID 838 850 BY SIMILARITY.
FT DISULFID 845 859 BY SIMILARITY.
FT DISULFID 861 873 BY SIMILARITY.
FT DISULFID 879 891 BY SIMILARITY.
FT DISULFID 885 900 BY SIMILARITY.
FT DISULFID 902 915 BY SIMILARITY.
FT DISULFID 921 933 BY SIMILARITY.
FT DISULFID 927 942 BY SIMILARITY.
FT DISULFID 944 957 BY SIMILARITY.
FT DISULFID 963 975 BY SIMILARITY.
FT DISULFID 970 984 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1109 1123 BY SIMILARITY.
FT DISULFID 1125 1138 BY SIMILARITY.
FT DISULFID 1144 1155 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1166 1179 BY SIMILARITY.
FT DISULFID 1298 1309 BY SIMILARITY.
FT DISULFID 1304 1318 BY SIMILARITY.
FT DISULFID 1320 1333 BY SIMILARITY.
FT DISULFID 1339 1354 BY SIMILARITY.
FT DISULFID 1349 1363 BY SIMILARITY.
FT DISULFID 1365 1378 BY SIMILARITY.
FT MOD.RES 647 647 HYDROXYLATION.
FT MOD.RES 810 810 HYDROXYLATION.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 1394 AA; 152791 MW; DFECA81A40B2C7D1 CRC64; /FTId-CAR_000184.

Query Match 22.48; Score 539; DB 1; Length 1394;
Best Local Similarity 29.68; Pred. No. 1e-33;
Matches 126; Conservative 51; Mismatches 126; Indels 122; Gaps 17;

QY 2 CTNGFDLDRSGGCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPVYRGPYSNPYST 61
DB 573 CYEGYRFSEQRCKVDIDECTQVQHLG-SQGRCENTGESFLCT----- 614
QY 62 PYSGRPAAPPLSAPRYPTISRLT-----CRFGYOMDESNOCCVDVD 104
DB 615 -CPAGMASEGTNCIDVDECLRPDVGEGHCVNTGAFRCCEYCDSDGYRMTQGRCECID 673
QY 105 ECATDHOCPNTOICINTEGGYTC-SCTDGYWLLGGCLDIDEC-RYGYCOQ-ILCANYPG 161
DB 674 ECLNPS-TC-PDQOCVNSPBGVCPCTEGFRGNNGOCLDVDECLFPVNCANGDCSNLEG 731
QY 162 SYSCGCPGFTLNEDGRSCQDVNECATENPCVQ----- 194
DB 732 SYMCSCHKGYTRPDHCHCRDIDECQOGLCNVNGCKNTGESFRCTCGGYQLSAKDDC 791
QY 195 -----TCVNTYGSFICRCDEYELBEDGVHCSMDDECSFSEFLQH-ECV 238
DB 792 EDIDECQHRHLCAHGOCRTNEGSFQCVCDGGRASGLDHCEDINECLEDRKSVCCORGDCI 851

QY 239 NQGTVECSQPGYLLDDNRSODINECEHNNHTCNLQGTQCYNLQGGFKCI----- 290
DB 852 NTAGSTICTCPDGF-QLDNKTQDINECEHPG-LCGPQGECLNTEGSFHCYCOGGSIS 909
QY 291 -DPIRCEPYLRIS-----DN-----RCMC-----PAENPGCRDOPFTLLYRDM 328
DB 910 ADRTCEIDECVNNVTCDSHGFCDNTRAGSFRLCYGQFQAPODQGCVD-----VNEC 963
QY 329 DVVSG 333
DB 964 ELLSG 968

Search completed: July 3, 2003, 18:23:14
Job time : 10.6843 secs

RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
 RX MEDLINE-91100426; PubMed-2269699;
 RA Argaves W.S., Tran H., Burgess W.H., Dickerson K.;
 RT "Fibulin is an extracellular matrix and plasma glycoprotein with
 RL repeated domain structure."
 RN J. Cell Biol. 111:3155-3164(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM D).
 RA Argaves S.;
 RN Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 567-703 FROM N.A. (ISOFORM C).
 RA Connor R.;
 RN Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 30-44.
 RX MEDLINE-8934537; PubMed-2527614;
 RA Argaves W.S., Dickerson K., Burgess W.H., Ruoslahti E.;
 RT "Fibulin, a novel protein that interacts with the fibronectin
 RL receptor beta subunit cytoplasmic domain."
 RN Cell 58:623-629(1989).
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS: A, B, C AND D (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-
 CC TERMINAL REGIONS.
 CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X53741; CAA37770.1; -
 DR EMBL: X53742; CAA37771.1; -
 DR EMBL: X53743; CAA37772.1; -
 DR EMBL: 001244; AAB17099.1; -
 DR EMBL: 295331; CAB62960.1; -
 DR PIR: A32826; A32826.
 DR PIR: A36346; A36346.
 DR PIR: B36346; B36346.
 DR PIR: C36346; C36346.
 DR HSP: P35555; 1EMN.
 DR Genew: HGNC:3600; FBLN1.
 DR MIM: 135820; -
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000361; EGF-like.
 DR InterPro: IPR001881; EGF-Ca.
 DR Pfam: PF00008; EGF_6.
 DR Pfam: PF01821; ANATO; 3.
 DR SMART: SM00104; ANATO; 3.
 DR SMART: SM00179; EGF_Ca; 7.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 DR PROSITE: PS01187; EGF_Ca; 8.
 KW Signal; Alternative splicing; Glycoprotein; Extracellular matrix;
 KW Repeat; EGF-like domain; Calcium-binding.
 FT CHAIN 1 29
 FT SIGNAL 1 29
 FT DOMAIN 36 76 ANAPHYLATOXIN-LIKE 1.
 FT DOMAIN 77 111 ANAPHYLATOXIN-LIKE 2.
 FT DOMAIN 112 144 ANAPHYLATOXIN-LIKE 3.
 FT DOMAIN 146 215 EGF-LIKE 1.
 FT DOMAIN 216 261 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 262 307 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

FT	DOMAIN	308	355	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	356	398	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	399	440	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	441	480	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	481	524	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	525	578	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	DISULFID	36	61	BY SIMILARITY.
FT	DISULFID	37	68	BY SIMILARITY.
FT	DISULFID	50	69	BY SIMILARITY.
FT	DISULFID	78	109	BY SIMILARITY.
FT	DISULFID	91	110	BY SIMILARITY.
FT	DISULFID	112	136	BY SIMILARITY.
FT	DISULFID	113	143	BY SIMILARITY.
FT	DISULFID	126	144	BY SIMILARITY.
FT	DISULFID	180	190	BY SIMILARITY.
FT	DISULFID	186	199	BY SIMILARITY.
FT	DISULFID	201	214	BY SIMILARITY.
FT	DISULFID	220	233	BY SIMILARITY.
FT	DISULFID	227	242	BY SIMILARITY.
FT	DISULFID	248	260	BY SIMILARITY.
FT	DISULFID	266	279	BY SIMILARITY.
FT	DISULFID	273	288	BY SIMILARITY.
FT	DISULFID	294	306	BY SIMILARITY.
FT	DISULFID	312	325	BY SIMILARITY.
FT	DISULFID	341	334	BY SIMILARITY.
FT	DISULFID	341	354	BY SIMILARITY.
FT	DISULFID	360	373	BY SIMILARITY.
FT	DISULFID	367	382	BY SIMILARITY.
FT	DISULFID	384	397	BY SIMILARITY.
FT	DISULFID	403	415	BY SIMILARITY.
FT	DISULFID	411	424	BY SIMILARITY.
FT	DISULFID	426	439	BY SIMILARITY.
FT	DISULFID	445	454	BY SIMILARITY.
FT	DISULFID	450	463	BY SIMILARITY.
FT	DISULFID	465	479	BY SIMILARITY.
FT	DISULFID	485	498	BY SIMILARITY.
FT	DISULFID	494	507	BY SIMILARITY.
FT	DISULFID	509	523	BY SIMILARITY.
FT	DISULFID	529	542	BY SIMILARITY.
FT	DISULFID	536	551	BY SIMILARITY.
FT	DISULFID	556	577	BY SIMILARITY.
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	535	535	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	539	539	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	567	703	MISSING (IN ISOFORM A).
FT	VARSPLIC	567	601	LOOEKDTVRCISCRPNVTCYFDPVHTISHYI -> QK SKKGRONTAGSSKEDCRVLPWKQGLDTHLDA (IN ISOFORM B).
FT	VARSPLIC	567	703	LOOEKDTVRCISCRPNVTCYFDPVHTISHYI EFTREPEILFLRAITPPHASPQANIIFDTEGMLRDSFDT KRYMDGMEVGVROVPIVGPFAVAKLENNYVGVGVSHR NVVNVRIIVSEVWF -> RCERLCHENRCSLTPLRIYX HLSFPNIOARPAVFERMGSSSAVPGSMOLALITGKNEGEF TTRKVSFHSGVAALTRPVEPRDLTLTKRMDLSRHGTSSSF VAKLFEVSAEL (IN ISOFORM C).
FT	CONFLICT	36	36	C -> S (IN REF. 4).
FT	CONFLICT	41	42	HR -> SH (IN REF. 4).
FT	SEQUENCE	703 AA;	72724 MW;	EPD88465BMA2D3A25 CRG64;

Query Match 28.8%; Score 694; DB 1; Length 703;
 Best Local Similarity 34.5%; Pred. No. 7.6e-46;
 Matches 159; Conservative 65; Mismatches 149; Indels 88; Gaps 21;

Oy	1	OCTNFFDLDROSGOGLDIDECTIEACGDMQVQNGVQNGYICIRPTNPNVRRPYNPS 60
Db	293	QCKSOFID-ALGNCIDINECLISAPCPDGTGHTINTGSIYC----- 334
Oy	61	TPYSGPYAAPPLPSAPVPTISRLPICRFGYOM-DESNQCVVDVDECATDSHOCNPQIC 119
Db	335	-----QKNVNP-----CGRGVHLINEGRTRCYVDVDCAPAPRCGSGHNC 373
Oy	120	INTEGGYTCSTGDIWTL--LBSGCLDIDEC-RY--GYCQQLCANPAGSYSCICNPGFTTN 174

CALCIUM DEPENDENT.
 -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
 -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH
 CC BASAL MEMBRANES AND OTHER CONNECTIVE TISSUES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.
 CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X82494; CAA57876.1; -
 DR HSSP: P00736; IAPQ.
 DR Genew: HGNC:3601; FBLN2.
 DR MIM: 135821; -
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_7.
 DR Pfam: PF01821; ANATO; 2.
 DR SMART: SM00104; ANATO; 3.
 DR SMART: SM00179; EGF_CA; 9.
 DR SMART: SM00001; EGF_Like; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 5.
 DR PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 DR PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 DR PROSITE: PS00022; EGF_1; FALSE_NEG.
 DR PROSITE: PS01186; EGF_2; 5.
 DR PROSITE: PS01187; EGF_CA; 9.
 DR Signal: Glycoprotein. Extracellular matrix; Plasma; EGF-like domain;
 DR Calcium-binding. Repeat.
 DR KMW
 DR SIGNAL
 FT CHAIN 1 27
 FT DOMAIN 28 1184
 FT DOMAIN 28 177
 FT DOMAIN 178 444
 FT DOMAIN 445 480
 FT DOMAIN 488 519
 FT DOMAIN 521 553
 FT DOMAIN 604 645
 FT DOMAIN 679 718
 FT DOMAIN 719 763
 FT DOMAIN 764 809
 FT DOMAIN 810 857
 FT DOMAIN 858 900
 FT DOMAIN 901 942
 FT DOMAIN 943 981
 FT DOMAIN 982 1024
 FT DOMAIN 1025 1069
 FT DOMAIN 1070 1184
 FT DISULFID 446 472
 FT DISULFID 446 479
 FT DISULFID 459 480
 FT DISULFID 489 518
 FT DISULFID 502 519
 FT DISULFID 521 545
 FT DISULFID 522 552
 FT DISULFID 535 553
 FT DISULFID 535 553
 FT DISULFID 608 620
 FT DISULFID 616 629
 FT DISULFID 631 644
 FT DISULFID 631 693
 FT DISULFID 689 702
 FT DISULFID 704 717
 FT DISULFID 723 736
 FT DISULFID 730 745
 FT DISULFID 751 762

FT DISULFID 768 781 BY SIMILARITY.
 FT DISULFID 775 790 BY SIMILARITY.
 FT DISULFID 796 808 BY SIMILARITY.
 FT DISULFID 814 827 BY SIMILARITY.
 FT DISULFID 821 836 BY SIMILARITY.
 FT DISULFID 843 856 BY SIMILARITY.
 FT DISULFID 862 875 BY SIMILARITY.
 FT DISULFID 869 884 BY SIMILARITY.
 FT DISULFID 886 899 BY SIMILARITY.
 FT DISULFID 905 917 BY SIMILARITY.
 FT DISULFID 913 926 BY SIMILARITY.
 FT DISULFID 928 941 BY SIMILARITY.
 FT DISULFID 947 956 BY SIMILARITY.
 FT DISULFID 952 965 BY SIMILARITY.
 FT DISULFID 967 980 BY SIMILARITY.
 FT DISULFID 986 998 BY SIMILARITY.
 FT DISULFID 994 1007 BY SIMILARITY.
 FT DISULFID 1009 1023 BY SIMILARITY.
 FT DISULFID 1029 1042 BY SIMILARITY.
 FT DISULFID 1036 1051 BY SIMILARITY.
 FT DISULFID 1056 1068 BY SIMILARITY.
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1184 AA; 126543 MW; CAA8490A55F9EC5D CRC64;
 Query Match 29.1%; Score 701.5; DB 1; Length 1184;
 Best Local Similarity 34.5%; Pred. No. 3.4e-46;
 Matches 142; Conservative 58; Mismatches 152; Indels 59; Gaps 11;
 QY 1 OCTNGFDLPDROSGCCLDIDECRTPEACRGDMKCVNONGYLCPRTNPKYRGYSNPYS 60
 DB 795 RCMDF-LADPEQNCVDINCTSLSEPCRPFSICINTVGSYTC----- 836
 QY 61 TPYSGFYPAAPPLSAPNRYTISPLICRGYQ-MDESNOCYVDYDECATDSHCNPTQIC 119
 DB 837 -----QRPNLICARVYHNSDQAKCVNDECTGVHRCGEQYV 875
 QY 120 INTEGGYTCSTGDYW--LLEGCCLDIDECRYG--CCQGLCANVPGSYSCNGFTLN 174
 DB 876 HNLPGSYRCDCRKGFGFQDARGRCIDVNECHASGRICQHNCTENTLSYRSCASGFLLA 935
 QY 175 EDGRSCQDVNECATENPCVQTCVNTVGSFICRCDPGYELBEDGVHCSMDSCS-FSEFLC 233
 DB 936 ADGKREDVNECEAQR-CSQECANIVGSYQCYCGQYQALADGHTCTDIDCAQAGILIC 994
 QY 234 QHECVQNPQGYFSCRP-PGTYLLDNRSCODINCEHRNHTCNLQCTCYNLQGGKCIDP 292
 DB 995 TFRCLVVPGSYQCAQCEQGYTMTANGRSCHDVADECALGTNHCSEAFCHNTIGSFRC-R 1053
 QY 293 IRCEEYLRISDRMCRCPAENPGCRD-----QFTLYRMDVYSGRSPADIPOMQAT 346
 DB 1054 FCECPNIVQVSKTKC-----ERTTCHDFLECONSPARITHVOLNPTGLVPAHFRIGPA 1109
 QY 347 TRYPGAYVYFQIKSGNEGREFYMRQTPISATLVMTPIKRPREIOLDLEM 397
 DB 1110 PAFGTPTALNITIKNDEGFFGRRLNAYGVYVLAQAVLEPRFADLVEM 1160
 RESULT 13
 FBLN_HUMAN STANDARD; PRT; 703 AA.
 AC P23142; P23143; P23144; P37888; Q90UR4;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrin-1 precursor.
 GN FBLN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_Taxid=9606;
 OX [1]

[illegible]

FT	CONFLICT	140	159	HSGRVAAHGHHVHSSCAC -> TVANISIMPRRPILIP
FT	CONFLICT	348	348	GF (IN REF. 2).
FT	CONFLICT	507	507	S -> L (IN REF. 2).
FT	CONFLICT	1102	1102	Q -> QO (IN REF. 2).
FT	CONFLICT	1221	1221	Q -> E (IN REF. 2).
SC	SEQUENCE	1221	AA; 131818	NW; 87DB2A10A8FDC45F CRC64;
Query Match				
Best Local Similarity		30.3%;	Score 729.5;	DB 1; Length 1221;
Matches 145;		Conservative 59;	Mismatches 148;	Indels 59; Gaps 11;
OY	1	QCTNGFDDLRSGGCLDIDECRTIPEACRGDMCMVONGYLICIPRTNPYRGFPYSNPS	60	
DB	832	RCMDGF-LQDEGNCVDINECTSLLEPCRSGFSCINTGVSYTC-----	873	
OY	61	TPYSGPYPAAPPLSAPNPYIPLISPLICRFGYOM-DESNOCVDYDECATSDHOCNPYIC	119	
DB	874	-----QRNPVLCGRGYHANEBSSECVDNECTGVHRCGEQGLC	912	
OY	120	INTEGGYTCSTGDYV-LLEGQCLDIDECRYG---YCOQLCANVPGSYSCTCNPGFTLN	174	
DB	913	YNLGSYKDCDKPFGORAPFOTCIDVNECMVSGRGLCHNCENTPSPGSRSCAAGFLTA	972	
OY	175	EDGNSCODVNECATENPCVOTCVNTYGSFICRDPGELEEDGYHSCDMDECS-FSEFLC	233	
DB	973	ADGHCHEBDECEFTRR-CSQECANIYGSCYCRGQYOLADHGHTCIDDECACGAGITLC	1031	
OY	234	QHECVNPGTYFGSCP-PGYILLDDNRSODINCEHNHNCNLQOTCYNLGGFKICDP	292	
DB	1032	TFRCVNVPGSTQCCAPEDGITYMMAANGRSCKLDIDECALGTNHCSEAFETHNLGGSKRL-R	1090	
OY	293	IRCEPFLYISDNRCMCPAENPGCD-----QFTLLYRMDVYSGRSVPADIFOMAT	346	
DB	1091	FDCPPNYRVAVSQTKC---ERTQODIECQTSFARIHYOINQOTGLVPAHIFRIGPA	1146	
OY	347	TRYGAYVYIPIKSGNREGEFMYOTGPIASLTAVTPRIKSPREIOLDLEM	397	
DB	1147	PAFAGDITSLITTKGNEEGYVYTRKLANAYTVASLQNSVLEPRFALDDEM	1197	
RESULT 11				
FBLL_MOUSE				
ID	FBLL_MOUSE	STANDARD:	PRT:	705 AA.
AC	Q08879; Q08878;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Fibulin-1 precursor (Basement-membrane protein 90) (BM-90).			
GN	FBNL1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=MK31.			
RX	MEDLINE=93358897; PubMed=8354280.			
RA	Pan T.-C., Kluge M., Zhang R.Z., Mayer U., Timpl R., Chu M.-L.,			
RT	"Sequence of extracellular mouse protein BM-90/fibulin and its			
RL	calcium-dependent binding to other basement-membrane ligands."			
RT	Eur. J. Biochem. 215:733-740(1993).			
CC	-1 SUBCELLULAR LOCATION: Secreted; extracellular matrix.			
CC	-1 ALTERNATIVE PRODUCTS: 4 ISOFORMS: A, B, C AND D (SHOWN HERE); ARE			
CC	PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER ONLY IN THEIR C-			
CC	TERMINAL REGIONS.			
CC	-1 SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.			
CC	-1 SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			

```

FT DISULFID 89 108 BY SIMILARITY.
FT DISULFID 110 134 BY SIMILARITY.
FT DISULFID 111 141 BY SIMILARITY.
FT DISULFID 124 142 BY SIMILARITY.
FT DISULFID 181 191 BY SIMILARITY.
FT DISULFID 187 200 BY SIMILARITY.
FT DISULFID 202 215 BY SIMILARITY.
FT DISULFID 221 234 BY SIMILARITY.
FT DISULFID 228 243 BY SIMILARITY.
FT DISULFID 249 261 BY SIMILARITY.
FT DISULFID 267 280 BY SIMILARITY.
FT DISULFID 274 289 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 313 326 BY SIMILARITY.
FT DISULFID 320 335 BY SIMILARITY.
FT DISULFID 342 355 BY SIMILARITY.
FT DISULFID 361 374 BY SIMILARITY.
FT DISULFID 368 383 BY SIMILARITY.
FT DISULFID 385 398 BY SIMILARITY.
FT DISULFID 404 416 BY SIMILARITY.
FT DISULFID 412 425 BY SIMILARITY.
FT DISULFID 427 440 BY SIMILARITY.
FT DISULFID 446 455 BY SIMILARITY.
FT DISULFID 451 464 BY SIMILARITY.
FT DISULFID 466 480 BY SIMILARITY.
FT DISULFID 486 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 524 BY SIMILARITY.
FT DISULFID 530 543 BY SIMILARITY.
FT DISULFID 537 552 BY SIMILARITY.
FT DISULFID 557 569 BY SIMILARITY.
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 638 638 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 684 AA: 75623 MW: 163807AD094739199 CMC64;

```

```

Query Match 32.0%; Score 769.5; DB 1; Length 684;
Best Local Similarity 36.5%; Pred. No. 1.3e-51;
Matches 159; Conservative 62; Mismatches 150; Indels 65; Gaps 15;

```

```

OY 1 QCNNGDLDNROSQCCLDIDCRITPEACRQDDMCNONGYLCTPRNPVYRGPSYNDYS 60
DB 294 QCNNGDLDNROSQCCLDIDCRITPEACRQDDMCNONGYLCTPRNPVYRGPSYNDYS 60
OY 61 TPYSGPYPAAPPLAPNPVTISAPLRCRGYQWDE-SNOCYVDDECATSDHQCNPFOIC 119
DB 341 TPYSGPYPAAPPLAPNPVTISAPLRCRGYQWDE-SNOCYVDDECATSDHQCNPFOIC 119
OY 120 INTEGGYTCSDTDGYW--LLEGQCLDIDECRY--GYCOQLCANVPGSYCTCNPGETLN 174
DB 375 INGPNGYRCCKSGSFDVLSFTCIDINECRARRGRCAHKCEMTPSYCTCMGFSLTS 434
OY 175 EGRSGQDVNECACTENPCVQVNTGSLFCRCRDPGLEEE-DGVHCSMDDECSF--SEF 231
DB 435 SGRSGEDNECES-SPCSOECANVGSYCYCRGQSLDIDIGISCDEDECALPTGGH 493
OY 232 LCOHECVNPGTGYFCSCP-PGYTILDDNRSQCCLDINECRHNRHNCNLQOTCVNLGGFRCI 290
DB 494 ICSFRCINIPSGQCCPSTGYGLARNAKNCODIDEVAETHNCSFNETCNIQGGFRL 553
OY 291 DTRCEEPRLRISDNFC-MCPA-ENPGCRDQPTIILYRMDVYSGRSVPADIFQMAQTR 348
DB 554 S-LCEPENYRKSGDTRCERLPCNENKCOQLPLRITYHLSFPTNIGVPTDIFMGPSNA 612
OY 349 YGAVYIPIKSGNENREFFYMGTPISALVMTPTPKGREGQLDLEMT-----YNTV 403
DB 613 VPGDXILSLISGNOGCFETTKKVNHSGLVVMOROTTEPRDLITOMOLTRHGTVNTF 672
OY 404 INFRGSSVIRLRIYVS 419
DB 673 I-----AKLFVFS 681

```

```

RESULT 10
FBI2_MOUSE
ID FBI2_MOUSE STANDARD: PRT: 1221 AA.
AC P37889: 09WU12;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibulin-2 precursor.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast;
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding.";
RL J. Cell Biol. 123:1269-1277(1993).
RN (2)
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99337686; PubMed=10406956;
RA Graessle S., Sticot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization.";
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC -1- CALCIUM DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and
CC 2/EGF3-less; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: X75285; CAA53040.1; -
DR EMBL: AF135253; AAD34456.1; -
DR EMBL: AF135239; AAD34456.1; JOINED.
DR EMBL: AF135240; AAD34456.1; JOINED.
DR EMBL: AF135241; AAD34456.1; JOINED.
DR EMBL: AF135242; AAD34456.1; JOINED.
DR EMBL: AF135243; AAD34456.1; JOINED.
DR EMBL: AF135244; AAD34456.1; JOINED.
DR EMBL: AF135245; AAD34456.1; JOINED.
DR EMBL: AF135246; AAD34456.1; JOINED.
DR EMBL: AF135247; AAD34456.1; JOINED.
DR EMBL: AF135248; AAD34456.1; JOINED.
DR EMBL: AF135249; AAD34456.1; JOINED.
DR EMBL: AF135250; AAD34456.1; JOINED.
DR EMBL: AF135251; AAD34456.1; JOINED.
DR EMBL: AF135252; AAD34456.1; JOINED.
DR PIR: A49457; A49457.
DR HSSP: P00736; IAPQ.
DR MGD: MGI:95488; Fbln2.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF; 6.

```


DR	PROSITE:	PS00022:	EGF_1:	FALSE_NEG.
DR	PROSITE:	PS01865:	EGF_2:	4
DR	PROSITE:	PS01877:	EGF_CA:	6
KW	Repeat:	EGF-like domain;	Calcium-binding; Glycoprotein; Signal;	
RN	Disease mutation:	Polymorphism; Alternative splicing.	POTENTIAL.	
FT	SIGNAL	1	17	
FT	CHAIN	18	493	
FT				EGF-CONTAINING FIBRIN-LIKE EXTRACELLULAR MATRIX PROTEIN 1.
FT	DOMAIN	26	71	EGF-Like 1, DIVERGENT.
FT		173	213	EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	214	253	EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	254	293	EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	294	333	EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	334	378	EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
FT	DISULFID	177	190	BY SIMILARITY.
FT	DISULFID	184	199	BY SIMILARITY.
FT	DISULFID	201	212	BY SIMILARITY.
FT	DISULFID	218	228	BY SIMILARITY.
FT	DISULFID	224	237	BY SIMILARITY.
FT	DISULFID	239	252	BY SIMILARITY.
FT	DISULFID	258	268	BY SIMILARITY.
FT	DISULFID	264	277	BY SIMILARITY.
FT	DISULFID	279	292	BY SIMILARITY.
FT	DISULFID	298	309	BY SIMILARITY.
FT	DISULFID	305	318	BY SIMILARITY.
FT	DISULFID	320	332	BY SIMILARITY.
FT	DISULFID	338	350	BY SIMILARITY.
FT	DISULFID	344 -	359	BY SIMILARITY.
FT	CARBOHYD	365	377	BY SIMILARITY.
FT	VARSPLIC	249	249	N-LINKED (GLUCNA...) (POTENTIAL).
FT	VARSPLIC	1	8	MISSING (IN ISOFORM 2).
FT	VARSPLIC	58	58	MISSING (IN ISOFORM 3).
FT	VARSPLIC	106	106	MISSING (IN ISOFORM 4).
FT	VARIANT	220	220	I -> F.
FT				/FTID=VAR_009512.
FT	VARIANT	345	345	R -> W (IN MULT).
FT				/FTID=VAR_009513.
SO	SEQUENCE	493 AA;	54640 MW;	128CASAEDI40DF414 CRC64;
	Query Match	Best Local Similarity	45.4%; Score 1093.5;	DB 1; Length 493;
	Matches 203;	Conservative 64;	Mismatches 152;	Indels 51; Gaps 7;
OY	1	OCTNGFDLDRSSGGCLDIDECRTIPPEARCRDMCNQNGYLCTIRTPMPV---	RGYSN	57
Db	28	QCTGDEMDPVROCKDIEDCDIYPADCAKGKMCVNHHGYLCLEPKTAQIIIVNNPOOE		87
OY	58	PY-----STPEYSGPPA-----	APPL	74
Db	88	TQPAEGTSGATTGVAASSMATISGLPGGGRVASAAAAGPEMGGRANNEYRRNPADPQ		147
OY	75	SAPNYPTITSLPLICRFGTQDNDSNOCVVDDECATDSHOCNPTQICINTEGGYTSCDTG		134
Db	148	RIPSNP--SHRIQCAAGEOSEHNVCODIDECITAGTHNCRADQVCINLRGSPACCCPGY		205
OY	135	WLLEGQCIDIDECRT-GYCQOLCANVPESYSTCCTCPGTLNEDHSQDVNCAENFCV		193
Db	206	QRKEQCVDDIECTIPPCCHORCVNTPSFGCCSPGRQLANNYTCDVINCDASNCA		265
OY	194	QTCVNTYGSIICRCDPGEYLEBEDGVHCSDMDECSSEFLCOHECVNQPGTYFCSCP	PXY	253
Db	266	QQCNINIGSFIQCNOGVELSSDLNCEIDIDECRTSYLQCYQCVNEBGRFSQCMCPGY		325
OY	254	LLDNRSGCODINEEHRNHVTNLQOCTYNLOGGFRCIDPIGCEBEYLAISDRKCAPEN		313
Db	326	VV--RSRFOCDINECEITNE-CREDEMKNYHGFRGCYPRNCCDYILTLPENRCVPSN		383
OY	314	PGCRDQFFILYRMADVVGSRVSPADIPOMQATTRPPAYAIIOIKSGNEGRFEFNROTG		373
Db	384	AMCREPOSIIYKKMISFSDRSVSDIFQIQATITTIANTIMTFRKSNEMNEGFLRQTS		443
OY	374	PISATLVMTIRIKGPREFIOLDLEMIVTVTNFRGSSVIRLRILIYSQYPF		423

	RESULT_9				
ID	FBL_CHICK	STANDARD;	PRT;	684 AA.	
AC	073775;				
DT	16-OCT-2001 (Rel. 4.0, Created)				
DT	16-OCT-2001 (Rel. 4.0, Last sequence update)				
DT	15-JUN-2002 (Rel. 4.1, Last annotation update)				
DE	Fibulin-1 precursor.				
GN	FBLN1.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
CC	Gallus.				
CX	NCBI_TaxID=9031;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99120531; PubMed=9923656;				
RA	Barch J.L., Argraues K.M., Roark E.F., Little C.D., Argraues W.S.;				
RT	"Identification of chicken and C. elegans fibulin-1 homologs and				
RL	characterization of the C. elegans fibulin-1 gene.";				
RL	Matrix Biol. 17:635-646(1998).				
CC	-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.				
CC	-1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.				
CC	-1- SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; AF051400; AAC05388.1; -				
DR	HSSP; P00742; IHCG.				
DR	InterPro; IPR000020; Anaphylatoxin.				
DR	InterPro; IPR000152; Asx_hydroxyl.				
DR	InterPro; IPR000561; EGF-like.				
DR	InterPro; IPR001881; EGF_Ca.				
DR	pfam; PF00008; EGF_6				
DR	pfam; PF01821; ANATO; 2.				
DR	SMART; SM0104; ANATO; 3.				
DR	SMART; SM00179; EGF_CA; 1.				
DR	SMART; SM00001; EGF_Like; 8.				
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.				
DR	PROSITE; PS01177; ANAPHYLATOXIN_1; 1.				
DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 2.				
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.				
DR	PROSITE; PS01186; EGF_2; 3.				
DR	PROSITE; PS01187; EGF_CA; 8.				
KW	Signal; Glycoprotein; Extracellular matrix; Repeat; EGF-like domain;				
KW	Calcium-binding.				
FT	SIGNAL	1..25	POTENTIAL.		
FT	CHAIN	26..684	FIBULIN-1.		
FT	DOMAIN	33..74	ANAPHYLATOXIN-Like 1.		
FT	DOMAIN	75..109	ANAPHYLATOXIN-Like 2.		
FT	DOMAIN	110..142	ANAPHYLATOXIN-Like 3.		
FT	DOMAIN	177..216	EGF-Like 1.		
FT	DOMAIN	217..262	EGF-Like 2; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	263..308	EGF-Like 3; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	309..356	EGF-Like 4; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	357..399	EGF-Like 5; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	400..441	EGF-Like 6; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	442..481	EGF-Like 7; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	482..525	EGF-Like 8; CALCIUM-BINDING (POTENTIAL).		
FT	DOMAIN	526..570	EGF-Like 9; CALCIUM-BINDING (POTENTIAL).		
FT	DISULFID	33..59	BY SIMILARITY.		
FT	DISULFID	34..66	BY SIMILARITY.		
FT	DISULFID	47..67	BY SIMILARITY.		
FT	DISULFID	76..107	BY SIMILARITY.		

```

DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 4.
DR PROSITE: PS01187; EGF_CA; 6.
KW Repeat, EGF-like domain; Calcium-binding; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 493
FT DOMAIN 26 71
FT DOMAIN 173 213
FT DOMAIN 214 253
FT DOMAIN 254 293
FT DOMAIN 294 333
FT DOMAIN 334 378
FT DISULFID 177 190
FT DISULFID 184 199
FT DISULFID 201 212
FT DISULFID 218 228
FT DISULFID 224 237
FT DISULFID 239 252
FT DISULFID 258 268
FT DISULFID 264 277
FT DISULFID 279 292
FT DISULFID 298 309
FT DISULFID 305 318
FT DISULFID 320 332
FT DISULFID 338 350
FT DISULFID 344 359
FT DISULFID 365 377
SQ SEQUENCE 493 AA; 54596 MW; 22DAFD70BACF1CA5 CRC64;

Query Match 46.1%; Score 1108.5; DB 1; Length 493;
Best Local Similarity 43.6%; Pred. No. 1.2e-77;
Matches 204; Conservative 68; Mismatches 149; Indels 47; Gaps 5;

QY 1 OCTGFGDLDROSGGCLDIDECRTIPACRGDMCMVQNGGYLCIPRTNPPYRGPPSNYS 60
DB 28 OCTGFGDLDROSGGCLDIDECRTIPACRGDMCMVQNGGYLCIPRTNPPYRGPPSNYS 87
QY 61 TPYS-----GPPAARPLSLAPNPT----- 81
DB 88 TPAAEASGAATGTAAASMATSGVTPGGGFTASATAVAGPEVOTGRNNEFVIRNPAPQ 147
QY 82 -----ISRLPFCRGYQMDENOCVAVDECATDSHCNPTQICINTEGGYSCSDGYWL 136
DB 148 RPSNPSPHRIQAGYEGSEHNVCODIDECSTGTHNCRIDQVCINLRGSGTCHLGEYOK 207
QY 137 LEGCCLLDICRY-GYCOQLCANVPGSYSCGTPFTLNEDGRSCODVNECATENPCVOT 195
DB 208 RGEQCVDIDECSPVPPYCHQGCYVTPGSPFCQCPGQFLAANNVTCVDINECDASNOCAQ 267
QY 196 CVNTGYSFRCRCDCGYLEDEGVHCSMDDECSFSELTQHECVNOCGYFECSPRGYIL 255
DB 268 CVNTLGSTCCNCGYELSDRLNCEIDECRTSYLCQYQCVNEGKSCMCPGGQYV 327
QY 256 DDNKSCDINECEHNHTCNLOQTCYNLGGFKCIDPICEBEPYRLISDNRCMCPAENPG 315
DB 328 -RSTCCDINCETTNE-CRDEKCMVNHGFRCPQNCOPRYVLTSENRCVCPYSNTM 365
QY 316 CRDOPFILLVDMQDVSGRVPADIFQMOATTRYPGAYYIFQIKSGNEGREFFMROGTPT 375
DB 386 CRDVPQIVYKVMNIRSDRSVPSDIFQJOATTIYANTINTFPRIKSGNENGEFFYLROTSPV 445
QY 376 SATLVNTRPIKGPRIQDLEMITVNTYINFRGSSVIRLRIYSGQPF 423
DB 446 SATLVNTRPIKGPRIQDLEMITVNTYINFRGSSVIRLRIYSGQPF 493

```

```

DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE EGF-containing fibulin-like extracellular matrix protein 1 precursor
DE (Fibulin-3) (FBL-3) (Fibrillin-like protein) (Extracellular protein
DE S1-5).
GN EFEMP1 OR FBLN3 OR FBNL
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND POSSIBLE ALTERNATIVE SPLICING.
RC MEDLINE=95097983; PubMed=7799918;
RA Lecka-Czenik B., Lumpkin C.K. Jr., Goldstein S.;
RT "An overexpressed gene transcript in senescent and quiescent human
RT fibroblasts encoding a novel protein in the epidermal growth factor-
RT like repeat family stimulates DNA synthesis.";
RL Mol. Cell. Biol. 15:120-128(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97001163; PubMed=8812496;
RA Ikegawa S., Toda T., Okui K., Nakamura Y.;
RT "Structure and chromosomal assignment of the human S1-5 gene (FBLN)
RT that is highly homologous to fibrillin.";
RL Genomics 35:590-592(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20068041; PubMed=10601734;
RA Giltay R., Timp R., Kostka G.;
RT "Sequence, recombinant expression and tissue localization of two novel
RT extracellular matrix proteins, fibulin-3 and fibulin-4.";
RL Matrix Biol. 18:469-480(1999).
RN [4]
RP VARIANT DHRD/MLVT TRP-345, AND VARIANT PHE-220.
RX MEDLINE=99295941; PubMed=10369267;
RA Stone E.M., Lotery A.J., Munier F.L., Heon E., Pignat B., Guymer R.H.,
RA Vandenburgh K., Cousin P., Nishimura D., Swiderski R.E., Silvestri G.,
RA Mackey D.A., Hagerman G.S., Bird A.C., Sheffield V.C.,
RA Schorderet D.F.;
RT "A single EFEMP1 mutation associated with both malattia Leventinese
RT and Doyme honeycomb retinal dystrophy.";
RL Nat. Genet. 22:199-202(1999).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; may
CC be produced by alternative splicing.
CC -1- DISEASE: DEFECTS IN EFEMP1 ARE A CAUSE OF DOYNE HONEYCOMB RETINAL
CC DYSTROPHY (DHRD) ALSO KNOWN AS MALATTIA LEVENTINESE (MLVT OR ML),
CC AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY YELLOW-WHITE
CC DEPOSITS KNOWN AS DRUSEN THAT ACCUMULATE BENEATH THE RETINAL
CC PIGMENT EPITHELIUM.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U03877; AAA65590.1; -.
DR HSSP: P35555; IEMN.
DR Genew: HGNC:3218; EFEMP1.
DR MIM: 601548; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR Pfam: PF00008; EGF; 3.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.

```

Ox	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Xl	[1]
Rn	NCBI_TaxId=10090;
Rp	SEQUENCE FROM N.A.
Rc	STRAIN=C57BL/6J;
Ra	MEDLINE=99308589; PubMed=10380882;
Rd	Gallagher W.M., Argentini M., Sierra V., Bracco L., Debussche L.,
Rt	Councillor E.;
Rf	"MBP1: a novel mutant p53-specific protein partner with oncogenic properties."
Rl	Oncogene 18:3608-3616(1999).
Cc	-1 SUBUNIT: BINDS PREFERENTIALLY TO P53 MUTANTS.
Cc	-1 SUBCELLULAR LOCATION: Secreted.
Cc	-1 SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
Cc	-----
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/or send an email to licenses@isb.slb.ch)
Cc	-----
Dl	EMBL; AF104223; AAD45219.1; .
Dl	HSSP; P00736; IAPQ.
Dl	MCD; MG1:1891209; Efemp2.
Dl	InterPro; IPRO00152; ASX_hydroxyL.
Dl	InterPro; IPRO00561; EGF-like.
Dl	InterPro; IPRO01881; EGF_Ca.
Dl	InterPro; IPRO01491; Thrmomodulin.
Dl	Pfam; Pf00008; EGF; 4.
Dl	PRINTS; PRO0907; THRMOMODULN.
Dl	SMART; SMO0179; EGF_CA; 4.
Dl	SMART; SMO0001; EGF_Like; 2.
Dl	PROSITE; PS00022; ASX_HYDROXYL; 4.
Dl	PROSITE; PS01186; EGF_1; FALSE_NEG.
Dl	PROSITE; PS01187; EGF_CA; 6.
Kw	Repeat; EGF-like domain; Calcium-binding; Glycoprotein; signal.
Ft	SIGNAL
Ft	CHAIN
Ft	1 25 POTENTIAL.
Ft	26 443 EGF-CONTAINING FIBRIN-LIKE EXTRACELLULAR MATRIX PROTEIN 2.
Ft	EGF-Like 1, DIVERGENT.
Ft	DOMAIN
Ft	123 163 EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
Ft	164 202 EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
Ft	203 242 EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
Ft	243 282 EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
Ft	283 328 EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
Ft	DISUFLID
Ft	127 140 BY SIMILARITY.
Ft	DISUFLID
Ft	134 149 BY SIMILARITY.
Ft	DISUFLID
Ft	151 162 BY SIMILARITY.
Ft	DISUFLID
Ft	168 177 BY SIMILARITY.
Ft	DISUFLID
Ft	173 186 BY SIMILARITY.
Ft	DISUFLID
Ft	188 201 BY SIMILARITY.
Ft	DISUFLID
Ft	207 217 BY SIMILARITY.
Ft	DISUFLID
Ft	213 226 BY SIMILARITY.
Ft	DISUFLID
Ft	228 241 BY SIMILARITY.
Ft	DISUFLID
Ft	247 258 BY SIMILARITY.
Ft	DISUFLID
Ft	254 267 BY SIMILARITY.
Ft	DISUFLID
Ft	269 281 BY SIMILARITY.
Ft	DISUFLID
Ft	287 300 BY SIMILARITY.
Ft	DISUFLID
Ft	294 309 BY SIMILARITY.
Ft	DISUFLID
Ft	315 327 BY SIMILARITY.
Ft	CASBOHD
Ft	198 198 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CASBOHD
Ft	394 394 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	SEQUENCE
Ft	443 AA; 49425 MW; 4969C032BA23DD8 CRC64;
Query March	52.9% Score 1272.5; DB 1; Length 443;
Best Local Similarity	52.0%; Pred. No. 3.2e+90;
Matches 220; Conservative	65; Mismatches 121; Indels 17; Gaps 3;
1	GCCTGGFFLDIROSGCCDLIDEERTIPDCRGDMCVVNGGYCLPTRTNPVRYCGPSYNDS 60

[illegible]

ID	FBL3_RAT	STANDARD	PRT	493 AA.
AC	035568			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	EGF-containing fibulin-like extracellular matrix protein 1 precursor			
DE	(Fibulin-3) (FBLN-3) (716 protein).			
GN	EFEMP1 OR FBLN3.			
OS	Rattus norvegicus (Rat).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID:10116;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=97415782; PubMed=9268694;			
RA	Ozeki T., Kondo K., Nakamura Y., Ichimiya S., Nakagawara A.,			
RA	Sakiyama S.;			
RT	"Interaction of D441, a DAN-binding protein, with the epidermal growth			
RT	factor-like protein, s(1-5).";			
RL	Biochem. Biophys. Res. Commun. 237:245-250(1997).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on the use			
CC	by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; D89730; BAA22265.1; .			
DR	HSSP; P35555; EMN.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	Pfam; PF00008; EGF_3.			
DR	SMART; SM00179; EGF_CA_4.			
DR	SMART; SM00001; EGF_like_2.			

Db 26 QCTNGFDLDROSGQCLDIDECRTIPEACRGDMCMVONGYLCIPRTNPVRYGPNPYS 85
 QY 61 TPYSGPAPAPPLSAPNPTISRLICRFYQMDENQCVYDDECATDSHOCNPTQICI 120
 Db 86 TTSYSGPAPAPPLSAPNPTISRLICRFYQMDENQCVYDDECATDSHOCNPTQICI 145
 QY 121 MNEGGYTCTDGYWLLGQCLDIDECRYGYCQQLCANVPSSYSTCNPFGTLNEDGRSC 180
 Db 146 MNEGGYTCTDGYWLLGQCLDIDECRYGYCQQLCANVPSSYSTCNPFGTLNEDGRSC 205
 QY 181 QDVNECATENPCVQVTCNTYGSFICRDPGYLEEDGVHCDMDCSSEFLCQHECVNQ 240
 Db 206 QDVNECATENPCVQVTCNTYGSFICRDPGYLEEDGVHCDMDCSSEFLCQHECVNQ 265
 QY 241 PGTYFSCPPGYILLDNRSCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 300
 Db 266 PGTYFSCPPGYILLDNRSCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 325
 QY 301 RISDNRCMCPAENPCRDQPTIYLRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 360
 Db 326 LIGENRCMCPAENPCRDQPTIYLRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 385
 QY 361 GNEGREYMRQGTPISTLWTRPIKGRPRETQIDLEMTVTNINFRSSVRLRLIYVSQ 420
 Db 386 GNEGREYMRQGTPISTLWTRPIKGRPRETQIDLEMTVTNINFRSSVRLRLIYVSQ 445
 QY 421 YPF 423
 Db 446 YPF 448

RESULT 4
 FBL4.CRIGR STANDARD: PRT: 443 AA.

AC 055058:
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE EGF-containing fibulin-like extracellular matrix protein 2 precursor
 GN (Fibulin-4) (FBLN-4) (H411 protein).
 GN EFEMP2 OR FBLN4.
 OS Crictetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Crictetidae;
 OC Crictetulus.
 NCBI_Taxid=10029;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT TISSUE=Ovary;
 RL Heine H., Delude R.L., Monks B., Golenbock D.T.;
 Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF046870; AAC03101.1;
 DR HSSP: P00736; IAPQ.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF_4.
 DR SMART: SM00179; EGF_CA; 4.
 DR SMART: SM00001; EGF_Like; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 4.
 DR PROSITE: PS00022; EGF_1; FALSE_NDG.
 DR PROSITE: PS01186; EGF_2; 4.

DR PROSITE: PS01187; EGF_CA; 6.
 KW Repeat: EGF-like domain; Calcium-binding; Glycoprotein; signal.
 FT SIGNL 1 25 POTENTIAL.
 FT CHAIN 26 443 EGF-CONTAINING FIBULIN-LIKE EXTRACELLULAR
 FT DOMAIN 36 81 MATRIX PROTEIN 2.
 FT DOMAIN 123 163 EGF-Like 1, DIVERGENT.
 FT DOMAIN 164 202 EGF-Like 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 203 242 EGF-Like 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 243 282 EGF-Like 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 283 328 EGF-Like 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 329 368 EGF-Like 6, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 127 140 BY SIMILARITY.
 FT DISULFID 134 149 BY SIMILARITY.
 FT DISULFID 151 162 BY SIMILARITY.
 FT DISULFID 168 177 BY SIMILARITY.
 FT DISULFID 173 186 BY SIMILARITY.
 FT DISULFID 188 201 BY SIMILARITY.
 FT DISULFID 207 217 BY SIMILARITY.
 FT DISULFID 213 226 BY SIMILARITY.
 FT DISULFID 228 241 BY SIMILARITY.
 FT DISULFID 247 258 BY SIMILARITY.
 FT DISULFID 254 267 BY SIMILARITY.
 FT DISULFID 269 281 BY SIMILARITY.
 FT DISULFID 287 300 BY SIMILARITY.
 FT DISULFID 294 309 BY SIMILARITY.
 FT DISULFID 315 327 BY SIMILARITY.
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 443 AA; 49432 MW; 0BCEP5D7329BE5F CRC64;

Query Match 53.08; Score 1276.5; DB 1; Length 443;
 Best Local Similarity 52.28; Pred. No. 1.6e-90;
 Matches 221; Conservative 64; Mismatches 121; Indels 17; Gaps 4;

QY 1 QCTNGFDLDROSGQCLDIDECRTIPEACRGDMCMVONGYLCIPRTNPVRYGPNPYS 60
 Db 38 ECTDGYEMDADSOHCRCVNECLTTPACRGDMKCNHGYGLCLPRAAIVNDLHG---- 93
 QY 61 TPYSGPAPAPPLSAPNPTISRLICRFYQMDENQCVYDDECATDSHOCNPTQICI 120
 Db 94 ---Egf-PPVYPPAQRHN-----PCPPYIEPDEQSCYDVDECAQALHDCRRSQDCH 141
 QY 121 MNEGGYTCTDGYWLLGQCLDIDECRYGYCQQLCANVPSSYSTCNPFGTLNEDGRSC 180
 Db 142 NLGSGYQCTCPDGYRKYRGPCEVDIDECRYRQCQRHCVNLPSFRCQCEPGFQGLGNNSC 201
 QY 181 QDVNECATENPCVQVTCNTYGSFICRDPGYLEEDGVHCDMDCSSEFLCQHECVNQ 240
 Db 202 VDVNECDMGAPCEQRCFNSYGTFLCRQNGYELHHDGFCSDIDECSSYLCQYRCVNE 261
 QY 241 PGTYFSCPPGYILLDNRSCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 300
 Db 262 PGTYFSCPPGYILLDNRSCQDINECEHRNHTCNLAQTCYNLAGFRCIDPICEEPTL 320
 QY 301 RISDNRCMCPAENPCRDQPTIYLRDMVYSGNSVPADIFOMATTRYPGAYYIFQIKS 360
 Db 321 QVSDNRCFCVSNPLCRQEPSSIVHRYMSITSESSVPADYFOIQTASYVPGAYNAFOIARA 380
 QY 361 GNEGREYMRQGTPISTLWTRPIKGRPRETQIDLEMTVTNINFRSSVRLRLIYVSQ 420
 Db 381 GNTQGDYTYRQINNVSAMLVARVGTGREYVLDLEMTVNSLMSYRASSVRLTLVFGA 440
 QY 421 YPF 423
 Db 441 YTF 443

RESULT 5
 FBL4.HUMAN STANDARD: PRT: 443 AA.
 ID FBL4.HUMAN
 AC 095967; 075967;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 136 153 BY SIMILARITY.
FT DISULFID 153 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 238 238 L -> P (IN REF. 2).
SO SEQUENCE 448 AA; 50160 MW; E6BC68F7BF14B714 CRC64;

Query Match
Best Local Similarity 95.7%; Score 2303; DB 1; Length 448;
Matches 401; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

OY 1 OCTNGFDLDRSGCCIDIDECRTIPACRGDMVCVNGGYLCIPRTNPPYRGYSNPS 60
DB 26 OCTNGFDLDRSGCCIDIDECRTIPACRGDMVCVNGGYLCIPRTNPPYRGYSNPS 85
OY 61 TPYSGPYPAAPPLSAPNYPTTISRPLICRGYOMDSNOCVDDECATDSHCNPTOICI 120
DB 86 TSYSGPYPAAPPLSAPNYPTTISRPLICRGYOMDSNOCVDDECATDSHCNPTOICI 145
OY 121 NTEGGYTCCTGGYWLLEGGCCIDIDECRGYCOOLCANVPGSSYCTCNPGFTLNEDGRSC 180
DB 146 NTEGGYTCCTGGYWLLEGGCCIDIDECRGYCOOLCANVPGSSYCTCNPGFTLNEDGRSC 205
OY 181 ODVNECATENPCVOTCVNTYGSFICRDPGELEBDSVHCSMDDECSFSEFLCOHECVNO 240
DB 206 ODVNECATENPCVOTCVNTYGSFICRDPGELEBDSVHCSMDDECSFSEFLCOHECVNO 265
OY 241 PGTYFCSGPPGYLLDDNRSQDINECEHNRHTCNLOQTCTYNLQGGFKCIDPIRCEBPYL 300
DB 266 PGTYFCSGPPGYLLDDNRSQDINECEHNRHTCTPLQCTYNNLQGGFKCIDPIRCEBPYL 325
OY 301 RISDNRCMCPAENPCGRDOPFTILYRDMDVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 360
DB 326 LIGDNRCMCPAENPCGRDOPFTILYRDMDVYSGRSVPADIFOMQATTRYPGAYYIFQIKS 385
OY 361 GNEGREFYRORTGPTISATLVMTKRGPREIOLDEMTYNTYINFGSSVIRLRIYVSQ 420
DB 386 GNEGREFYRORTGPTISATLVMTKRGPREIOLDEMTYNTYINFGSSVIRLRIYVSQ 445
OY 421 YPF 423
DB 446 YPF 448

RESULT 3
FBL5_MOUSE STANDARD; PRT; 448 AA.
ID FBL5_MOUSE
AC O9MVH9;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DE F1bulin-5 precursor (F1bul-5) (Developmental arteries and neural crest
DE EGF-like protein) (Dance).
GN FBLN5 OR DANCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357779; PubMed=10428923;
RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Taniwaki M.,
RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;
RT "DANCE", a novel secreted RGD protein expressed in developing,
RT atherosclerotic, and balloon-injured arteries.";
RL J. Biol. Chem. 274:22476-22483(1999).
CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
CC DEVELOPMENT AND REMODELING.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF112151; AM041767.1; -
DR HSP; P00736; IAP0.
DR GMD; MG1:134691; Fbln5.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR Pfam; PF00008; EGF; 4.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_Like; 2.
DR PROSITE; PS00010; Asx_Hydroxyl; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS01187; EGF_CA; 6.
KM Repeat; EGF-like domain; Calcium-binding; Glycoprotein; signal.
FT SIGNAL 1 23
FT CHAIN 24 448
FT DOMAIN 24 69 EGF-LIKE 1, DIVERGENT.
FT DOMAIN 167 177 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 168 206 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 207 246 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 247 287 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 288 333 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 131 144 BY SIMILARITY.
FT DISULFID 138 153 BY SIMILARITY.
FT DISULFID 155 166 BY SIMILARITY.
FT DISULFID 172 181 BY SIMILARITY.
FT DISULFID 177 190 BY SIMILARITY.
FT DISULFID 192 205 BY SIMILARITY.
FT DISULFID 211 221 BY SIMILARITY.
FT DISULFID 217 230 BY SIMILARITY.
FT DISULFID 232 245 BY SIMILARITY.
FT DISULFID 251 262 BY SIMILARITY.
FT DISULFID 258 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 320 332 BY SIMILARITY.
FT CARBOHYD 283 283 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 448 AA; 50193 MW; F15CC70CCEBFD97 CRC64;

Query Match
Best Local Similarity 95.6%; Score 2302; DB 1; Length 448;
Matches 401; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 1 OCTNGFDLDRSGCCIDIDECRTIPACRGDMVCVNGGYLCIPRTNPPYRGYSNPS 60

```

DR Genew; HGNC:3602; FBIN5.
 DR MIM; 604580; -
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam: PF00008; EGF; 4.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
 FT CHAIN 1 23
 FT SIGNAL 1 23
 FT DOMAIN 24 448
 FT DOMAIN 127 167
 FT DOMAIN 168 206
 FT DOMAIN 207 246
 FT DOMAIN 247 287
 FT DOMAIN 288 333
 FT SITE 54 56
 FT DISULFID 131 144
 FT DISULFID 138 153
 FT DISULFID 155 166
 FT DISULFID 172 181
 FT DISULFID 177 190
 FT DISULFID 192 205
 FT DISULFID 211 221
 FT DISULFID 217 230
 FT DISULFID 232 245
 FT DISULFID 251 262
 FT DISULFID 258 271
 FT DISULFID 273 286
 FT DISULFID 299 314
 FT DISULFID 320 332
 FT CARBOHYD 283 283
 FT CARBOHYD 296 296
 FT CONFLICT 69 70
 FT CONFLICT 147 148
 FT CONFLICT 147 148
 SQ SEQUENCE 448 AA; 50180 MW; 19FCAS1FDA328003 CRC64;

Query Match 100.0%; Score 2407; DB 1; Length 448;
 Best Local Similarity 100.0%; Pred. No. 7.5e-177;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCVNONGYLCIPRTNVPYRGPSNPYS 60
 26 QCTNGFPLDRSGGCLDIDECRTPEACRGDMCVNONGYLCIPRTNVPYRGPSNPYS 85
 61 TPYSGPAPAAPPTISAPNYPTISRLPILCFRGQMDSESNOCVYDECATDSHOCNPOICI 120
 86 TPYSGPAPAAPPTISAPNYPTISRLPILCFRGQMDSESNOCVYDECATDSHOCNPOICI 145
 121 NTEGGYTCSCDGYWLLGCGCLDIDECRYGYCQOLCANVGSYSCTCNPGFTLNEDGRSC 180
 146 NTEGGYTCSCDGYWLLGCGCLDIDECRYGYCQOLCANVGSYSCTCNPGFTLNEDGRSC 205
 181 QDVNECATENPCVOTCVNTYSSFTICRCPGYLEEDGVCHSCMDECSSEFELCOHCVCNO 240
 206 QDVNECATENPCVOTCVNTYSSFTICRCPGYLEEDGVCHSCMDECSSEFELCOHCVCNO 265
 241 PGTYFGSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGGFKCIDPRCEPEYTL 300
 266 PGTYFGSCPPIYLLDNRSCODINECHRNHTCNLQOTCYNLGGGFKCIDPRCEPEYTL 325
 301 RISDNRCMCPAENGCGRDQPTILYRDMDVYSGRSVPADIFQMATRTTTPGAYYIFQIKS 360
 326 RISDNRCMCPAENGCGRDQPTILYRDMDVYSGRSVPADIFQMATRTTTPGAYYIFQIKS 385
 361 GNEGREYMRQGTISATLVWTRPIKGPKEIOLDLEMTVNTVINFRGSSVIRLRIVYSO 420

DB 386 GNEGREYMRQGTISATLVWTRPIKGPKEIOLDLEMTVNTVINFRGSSVIRLRIVYSO 445
 OY 421 YPF 423
 DB 446 YPF 448

RESULT 2
 FBLS_RAT
 ID FBLS_RAT STANDARD: PRT: 448 AA.
 AC Q9WVH8; Q9R284;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Fblulin-5 precursor (FBLS-5) (Developmental arteries and neural crest
 DE EGF-like protein) (Dance) (Embryonic vascular EGF repeat-containing
 DE protein) (EVEC).
 GN FBLS OR DANCE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99357779; PubMed=10428823;
 RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,
 RA Furukawa Y., Kobike K., Tashiro K., Lu Z., Andon N.L., Schaub R.,
 RA Matsunori A., Sasayama S., Chien K.R., Honjo T.;
 RT "DANCE, a novel secreted RGD protein expressed in developing,
 RT atherosclerotic, and balloon-injured arteries.";
 RL J. Biol. Chem. 274:22476-22483(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99278197; PubMed=10347091;
 RA Kwee R.C., Richardson J.A., Milano J.M., Olson E.N.;
 RT "EVEC, a novel epidermal growth factor-like repeat-containing protein
 RT upregulated in embryonic and diseased adult vasculature.";
 RL Circ. Res. 84:1166-1176(1999).
 CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH
 CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR
 CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR
 CC DEVELOPMENT AND REMODELING.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF112153; AAD41769.1; -;
 DR EMBL; AF137350; AAD25101.1; -;
 DR HSSP; P00736; IAPQ.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-Like.
 DR InterPro: IPR001881; EGF_Ca.
 DR Pfam; PF00008; EGF; 4.
 DR SMART; SM00179; EGF_CA; 4.
 DR SMART; SM00001; EGF_Like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 4.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS01187; EGF_CA; 6.
 KW Repeat; EGF-like domain; Calcium-binding; glycoprotein; signal.
 FT CHAIN 1 23
 FT SIGNAL 1 23
 FT DOMAIN 24 448
 FT DOMAIN 127 167
 FT DOMAIN 168 206
 FT DOMAIN 207 246

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:40:07 ; Search time 10.6843 Seconds
(without alignments)
1642.086 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFDLDROSGOCLDIDE.....INFRGSSVRLRIYVSQIPF 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2407	100.0	448	1	FBL5_HUMAN
2	2303	95.7	448	1	FBL5_RAT
3	2302	95.6	448	1	FBL5_MOUSE
4	1276.5	53.0	443	1	FBL4_CRIGR
5	1276.5	53.0	443	1	FBL4_HUMAN
6	1272.5	52.9	443	1	FBL4_MOUSE
7	1108.5	46.1	493	1	FBL3_RAT
8	1093.5	45.4	493	1	FBL3_HUMAN
9	769.5	32.0	684	1	FBL1_CHICK
10	729.5	30.3	1221	1	FBL2_MOUSE
11	702	29.2	705	1	FBL1_MOUSE
12	701.5	29.1	1184	1	FBL2_HUMAN
13	694	28.8	703	1	FBL1_HUMAN
14	596	24.8	712	1	FBL1_CAEEL
15	539	22.4	1394	1	LTBL_HUMAN
16	539	22.4	1595	1	LTBL_HUMAN
17	523.5	21.7	2871	1	FBN1_HUMAN
18	522.5	21.7	2871	1	FBN1_BOVIN
19	518	21.5	1712	1	LTBL_RAT
20	516.5	21.5	2871	1	FBN1_PIG
21	509.5	21.2	2871	1	FBN1_MOUSE
22	509.5	21.2	2907	1	FBN2_MOUSE
23	508.5	21.1	2911	1	FBN2_HUMAN
24	455	18.9	956	1	MTN2_HUMAN
25	455	18.9	956	1	MTN2_MOUSE
26	431	17.9	956	1	MTN2_MOUSE
27	392	16.3	886	1	EMRI_MOUSE
28	375	15.6	2470	1	NTC2_MOUSE
29	372	15.5	810	1	NEIL_HUMAN
30	372	15.5	810	1	NEIL_MOUSE
31	369	15.3	2471	1	NTC2_HUMAN
32	367	15.2	2471	1	NTC2_HUMAN
33	364	15.1	810	1	NEIL_RAT

34	361.5	15.0	816	1	NEIL_HUMAN	Q99435 homo sapien
35	357.5	14.9	816	1	NEIL_RAT	O62918 rattus norv
36	355.5	14.8	816	1	NEIL_CHICK	O90827 gallus gall
37	353.5	14.7	1964	1	NTC4_MOUSE	P31656 mus musculu
38	351.5	14.6	2437	1	NTC1_BRARE	P46530 diacydantio
39	350.5	14.6	652	1	CD93_HUMAN	Q99433 homo sapien
40	348.5	14.5	2703	1	NOTC_DROME	P07207 drosophila
41	347.5	14.4	644	1	CD93_MOUSE	O69103 mus musculu
42	345.5	14.4	2556	1	NTC1_HUMAN	O89103 mus musculu
43	340	14.1	2524	1	NOTC_XENTLA	P46531 homo sapien
44	337	14.0	835	1	CD97_HUMAN	P21783 xenopus lae
45	335.5	13.9	2531	1	NTC1_RAT	P48960 homo sapien
						Q07008 rattus norv

ALIGNMENTS

RESULT 1
FBL5_HUMAN STANDARD; PRT; 448 AA.

AC Q99435: 075966; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fblulin-5 precursor (FBL-5) (Developmental arteries and neural crest

DE EGF-like protein) (Dance) (Urine p50 protein) (UP50).

DE EGF-LIKE OR DANCE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.

OX NCBI_Taxid:9606;

RP SEQUENCE FROM N. A.

RA TISSUE-Melanoma;

RL Kostka G., (MAR-1999) to the EMBL/Genbank/DBJ databases.

RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.

RL [2]

RP SEQUENCE FROM N. A.

RX MEDLINE:99357779; PubMed:10428823;

RA Nakamura T., Ruiz-Lozano P., Lindner V., Yabe D., Tanikawa M.,

RA Furukawa Y., Kobuke K., Tashiro K., Lu Z., Andon N.L., Schaub R.,

RA Matsumori A., Sasayama S., Chien K.R., Honjo T.;

RT "DANCE, a novel secreted RGD protein expressed in developing,

RT atherosclerotic, and balloon-injured arteries."

RL J. Biol. Chem. 274:22476-22483(1999).

RL [3]

RP SEQUENCE FROM N. A.

RC TISSUE-Urine;

RA Zemel R., Sholto O., Shaul Y.;

RA Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: PROMOTES ADHESION OF ENDOTHELIAL CELLS THROUGH

CC INTERACTION OF INTEGRINS AND THE RGD MOTIF. COULD BE A VASCULAR

CC LIGAND FOR INTEGRIN RECEPTORS AND MAY PLAY A ROLE IN VASCULAR

CC DEVELOPMENT AND REMODELING.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN HEART, OVARY, AND

CC COLON BUT ALSO IN KIDNEY, PANCREAS, TESTIS, LUNG, AND PLACENTA.

CC NOT DETECTABLE IN BRAIN, LIVER, THYMUS, PROSTATE, OR PERIPHERAL

CC BLOOD LEUCOCYTES.

CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC EMBL: AJ133490; CAB38568.1; -

CC EMBL: AF112152; AAD41768.1; -

CC EMBL: AF093118; AAC62107.1; -

CC HSSP: P00736; IAP0.

A:Cross-references: GB:M34057; NID:9339547; PTDN:AAA61160.1; PID:9339548
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing
 F:750-791/Domain: EGF homology <EGF>

Query Match 22.4%; Score 539; DB 2; Length 1394;
 Best Local Similarity 29.6%; Pred. No. 4,6e-28;
 Matches 126; Conservative 51; Mismatches 126; Indels 122; Gaps 17;

QY 2 CTGPFDRSGGCLDDECRITPEACRGMCMVONGYLCIPRTNPNVNGPSNPYST 61
 DB 573 CTEGYRFEQQRQCVDDICTQVQHLCSGRCENTEGSFLLCI----- 614
 QY 62 PYSGPYAAPAPISAPNYPTISRPLI-----CRFGYOMDESNCVVD 104
 DB 615 -CPAGFMASBEETNCIDVDDECLRPDVGEGHCVTVGAFREYCDSGIRMRGRCEID 673
 QY 105 ECATDSHQCNPTQICINTEGGYTC-SCTDGYWLEGGCLDDEC-RVGYCOO-ICANYPG 161
 DB 674 ECLNPS-TC-PDEQCVNPSGSCVPCFTEGFRGNNGCLDDECLPNCVANGCSNLEG 731
 QY 162 SYSCNCPFTLNEDGRSCDVNDECATENECVQ----- 194
 DB 732 STWCSCHKGTTRPDHRCRDIDECQCGNLCVNGQCKNTGSRCTCGGYQLSMAKQDC 791
 QY 195 -----TCVNTYGSFICRDPGYELEEDGVHCSDMDECSFSELCOH-ECV 238
 DB 792 EDIDECQHRLCAHGCRNTEGSEFCQCDGYSRSGLDHCEIDNECLDKRSVCORGDCT 851
 QY 239 NQPTGFCSCPPGYILLDNRSQDINECEHRNHTCNLOTCYNLQGGFKCI----- 290
 DB 852 NTAQSYDCCTCPDPE-GLDPMKTCODINECEHPG-TCSPQEGCLNTEGSHFCVCGQGFIS 909
 QY 291 -DPIRCEEPYLRIS-----DN-----RCMC-----PANNPCGRDPFTILLYDM 328
 DB 910 AGRCEDEDDECYNNTVCSHGRCNDNTAGSFRCCLCYGFOAPDQGGCVD-----VNEC 963
 QY 329 DVGSG 333
 DB 964 ERLSG 968

RESULT 14

A47221
 fibrillin 1 precursor - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 02-Jun-1995 #sequence, revision 25-Apr-1997 #text, change 02-Aug-2002
 C:Accession: A47221; S15355; S17064; I59574; S17062; S62111; A34198
 R:Corson, G.M.; Chabberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.
 Genomics 17, 476-484, 1993
 A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure
 A:Reference number: A47221; MUID:94010947; PMID:7691719
 A:Accession: A47221
 A:Molecule type: mRNA
 A:Residues: 1-337, 'T', 339-1029 <COR>
 A:Cross-references: GB:X63556
 R:Perle, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Panglman, T.; Bonad
 Hum. Mol. Genet. 2, 961-968, 1993
 A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene F
 A:Reference number: I54355; MUID:93372860; PMID:8364578
 A:Accession: I54355
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 132-3002 <PER>
 A:Cross-references: GB:11923; NID:9306745; PTDN:AAB02036.1; PID:9306746
 R:Maalen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.
 Nature 352, 334-337, 1991
 A:Title: Partial sequence of a candidate gene for the Marfan syndrome.
 A:Reference number: S17064; MUID:91304568; PMID:1852207
 A:Accession: S17064
 A:Molecule type: mRNA
 A:Residues: 1030-3002 <MAS>
 A:Cross-references: EMBL:X63556

R:Dietz, H.C.; Valle, D.; Franccomano, C.A.; Kendzior, R.J.
 Science 259, 680-683, 1993
 A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.
 A:Reference number: I59574; MUID:93157831; PMID:8430317
 A:Accession: I59574
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 2217-2288, 'T', 2290-2325 <RES>
 A:Cross-references: GB:S54426; NID:9264860; PTDN:AAB25244.1; PID:9264861
 R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras,
 Nature 352, 330-334, 1991
 A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff
 A:Reference number: S17062; MUID:91304567; PMID:1852206
 A:Accession: S17062
 A:Molecule type: mRNA
 A:Residues: 'VLTVPVPIFLSYNKM', 944-1444 <LEE1>
 A:Cross-references: EMBL:X62008; NID:931398; PTDN:CAB56534.1; PID:95924015
 A:Accession: S62111
 A:Molecule type: protein
 A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>
 R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.
 J. Biol. Chem. 264, 21881-21885, 1989
 A:Title: Connective tissue microfibrils. Isolation and characterization of three larg
 A:Reference number: A34198; MUID:90078246; PMID:2512293
 A:Accession: A34198
 A:Molecule type: protein
 A:Residues: 565-575; 1890-1892, 'T', 1894-1900 <MAD>
 C:Comment: Fibrillin is a major component of elastin-associated microfibrils.
 A:Gene: GDB:FBNI
 A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700
 A:Map position: 15q21.1-15q21.1
 A:Antons: 2236/1; 2258/1; 2287/1
 C:Superfamily: fibrillin 1; EGF homology
 C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein
 F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predict
 F:1332-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MARC>
 F:1332-1367/Domain: EGF homology <EGF>
 F:1457-1492/Domain: EGF homology <EGF2>
 F:2262-2295/Domain: EGF homology <EGF1>

Query Match 21.7%; Score 523.5; DB 2; Length 3002;
 Best Local Similarity 30.3%; Pred. No. 9.5e-27;
 Matches 122; Conservative 47; Mismatches 120; Indels 113; Gaps 18;

QY 1 OCTNGFDLDRSGGCLDDECRITPEACRGMCMVONGYLCIPRTNPNVNGPSNPYS 60
 DB 1185 RCDSEFALDSEBRNCTDIDECRISPLIC-GRGQCVNTPGDFEC-KCDEGYESGF----- 1236
 QY 61 TPYSGPYAAPAPISAPNYPTISRPLI-----GYOMDES-NO 99
 DB 1237 -----MMKNKCMIDDECORDPULCGVCHNTEGSRCECPGHOLSPNISA 1283
 QY 100 CVDVDECATDSHQCNPTQICINTEGGYTCSDGYWLEGGQ-CLDIDECRY--GYCOOL 155
 DB 1284 CIDINECELSAHLIC-PNGRCVNLIGKYQACANGYSTPRLFCVDDIDESINNCGETF 1342
 QY 156 CANVGSYSCTGNPFTLNEDGRSCDVNDECATENP----- 191
 DB 1343 CTNSGSEYECSCQPGFALPMDORSCTDIDEC-EDNNICDGGQCTNIPGEYRCICYDGF 1401
 QY 192 -----CVQ-----TCVNTYGSFICRDPGYELEEDGVHCSDMDECSFSE 230
 DB 1402 ASEDKMTVDVNECLNPTQICINTEGGYTCSDGYWLEGGQ-CLDIDECRY--GYCOOL 1461
 QY 231 FLC-OHE-CVNOPGTGFCSCPPGYILLDNRSQDINECEHRNHTCNLOTCYNLQGGFK 288
 DB 1462 HNCGHAVCTMTAGSFKSCSPGWT--GGINCTDIDECNSGTHMCSQAHDCNTMGSYR 1519
 QY 289 CIDPICEEPY-----RISDNRCMCAENPCGRDOP 320
 DB 1520 CL-----CKEGYTGDFCTDIDECSENLNLC--GNGQCLINAP 1555

submitted to the EMBL Data Library, June 1998

A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character

A:Reference number: T22337

A:Accession: T43210

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-569 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: Intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 24.6%; Score 592.5; DB 2; Length 589;

Best Local Similarity 27.4%; Pred. No. 6.3e-32;

Matches 156; Conservative 52; Mismatches 155; Indels 207; Gaps 20;

2 CINGFDLRSGGCLD-IDECRTIPEACRGDMCMVQNGYLCPRTNPTVYRGPISNYS 60

79 CRSGFDLPDGMACVCHIDECATLMDCLSRQCLNPPGSKCI----- 122

61 TPYSGYPAAPLAPLAPNPTISRLICRFGYOMD-ESNOCVDYDECATDSHOCNPTQIC 119

123 -----RTLSGCGIYAMDSETERCRDVECLGSHDCGPLYC 159

120 INTEGTYC-----SCIDGYMLLEGQCLDIDECRYG----- 151

160 RMTQGYRCDARKKCGDELQNPMTGECTSTICPNGYVPKMGCMNDIDECVGHNCGAGE 219

152 -----CGQ---LCA-----NVPGSYCTCNG 170

220 CVNTPGSRQOGKGNLCAMHYEVNGATGCEDEVNCOQGVGSMECIMLPETVCKCKGPG 279

171 FTLE-----DGRSCODVNEC 186

280 YENDAKKCEDVEDCKIKFAGHYCDLSAECINTIGSEPCCKKPGFQLASDGRKCDVNEC 339

187 AFE-NPCVOTCVNTYSGFICRDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNOP 241

340 TGIACCEKCVNIPGSGYQICIDRGFALGPDGKCEDIDECSIMAGSGNDLGMGCMWK 399

242 GMYFSGCPGYILLDNRSQDINECEHNTNMLQOTCVNLOGGFKCIDPIRCEPIYR 301

400 GSTLCCCPGYKIQPDGRTVDVDECA-MGECAGSDKCVNTLGSFK-HSIDCPNTYIH 457

302 ISDNR-----C-MCPAENPGC-RDOPFTILYRDMVYSGRSV--PADI----- 340

458 DSLNKQIADGVGICIVCSFEDTECLGNHTRVLYKQRAVPSLKTIIISPIEVSRTVING 517

341 --FOMQATRYPGAYYIFQIKSGNEGREFTWROTGPISATLWTRPIKGPRIQDLEMT 398

518 VFESVDYNDLYVQGRHRIVQERNIG-----IVQLVKPISGP-----TVEYI 559

399 TWTVTINFRGSSYIR-----LRIYVSQYRF 423

560 KVNIMTKSRGCVTLAENEALIEISVSKYP 589

Db

RESULT 12

T22793

hypothetical protein F56H1.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Aug-2002

C:Accession: T22793; T24489

R:Wilkinson, J.

submitted to the EMBL Data Library, January 1996

A:Reference number: Z19616

A:Accession: T22793

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <MIL>

A:Cross-references: EMBL:Z68749; PIDN:CAA92962.1; GSPDB:GNO0022; CESP:F56H1.1

A:Experimental source: clone F56H1

R:Lloyd, C.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19897

A:Accession: T24489

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-798 <MIL>

A:Cross-references: EMBL:Z68219; PIDN:CAA92483.1; GSPDB:GNO0022; CESP:F56H1.1

A:Experimental source: clone T05A1

C:Genetics:

A:Gene: CESP:F56H1.1

A:Map position: 4

A:Introns: 14/1; 92/3; 144/1; 195/1; 281/1; 325/2; 371/2; 390/1; 437/1; 471/3; 498/3;

C:Superfamily: fibulin-1; EGF homology

Query Match 23.8%; Score 574; DB 2; Length 798;

Best Local Similarity 26.8%; Pred. No. 1.4e-30;

Matches 139; Conservative 49; Mismatches 172; Indels 214; Gaps 22;

2 CINGFDLRSGGCLD-IDECRTIPEACRGDMCMVQNGYLCPRTNPTVYRGPISNYS 51

180 CRSGFDLPDGMACVCHIDECATLMDCLSRQCLNPPGSKCI-----IP---RTNPTV 239

52 RGP---YSNPTSPYSGPYPAAP----- 73

240 NABRRMRDDYSR--AGEYRASAQANTEFGCPMGMLFQHGHCVIDECATLMDCLSR 297

74 -LSAPNPTISRLICRFGYOMD-ESNOCV-----DVDECATDSHOC 113

298 CLNTPGSRKCIKRTLSGCTGYAMDSETERNNCFLLIINNTFNCKTFVEVDVDCNLGSHDC 357

114 NPTQICINTEGTYCS----- 129

358 GPLQCKNTQGSYRCDARKKCGDELQNPMTGEYIDECVTHNCAGAGECVNTPGSRQ 417

130 -----CIDGYML--LEGQCLDIDECRYGCOOL-CANVGSYCTCNPFTLE----- 175

418 KGNLCAMHYEVNGATGCEDEVNCOQGVGSMECIMLPETVCKCKGPGYEFNDARKCED 477

176 -----DGRSCODVNECATE-NPCVQYCV 197

478 VDECIFKAGHYCDLSAECINTIGSEPCCKKPGFQLASDGRKCDVNECTTGIACDEKCV 537

198 NTYSGFICRDPGYLEEDGVHCSMDDECSF-----SEFLQHECVNOPGYFCSPGYI 253

538 NIGSYQICIDRGFALGPDGKCEDIDECSIMAGSGNDLGMGCMWKINTKSGYLQCPGPK 597

254 LDDNSQDINECEHNTNMLQOTCVNLOGGFKCIDPIRCEPIYRISDNR----- 306

598 IOPDGRCTVDVDECA-MGECAGSDKCVNTLGSFK-HSIDCPNTYIHDSLKNKRCNOP 655

307 --CMCPAENPGCRDOPFTILYRDMVYSGRSV-----ADI---FOM 343

656 SAGCLPBE---CSKVPLEFLTYQFISL--ARAVPISSHRPATILFKVSPANMADTEVNEFL 710

344 QATTRYPGAYYIFQIKSGNEGREFTWROTGPISATLWTRPIKGPRIQDLEMT 395

711 QKTTIVGADNVLPALIRAN-----FLQKGEKRSAAVTVTLRDSLDPQTVKQL 759

Db

RESULT 13

A35626

transforming growth factor beta-1-binding protein - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 11-Jan-2000

C:Accession: A35626

R:Kanzaki, T.; Olofsson, A.; Momen, A.; Wernstedt, C.; Hellman, U.; Miyazono, K.; Cla

Cell 61, 1051-1061, 1990

A:Title: TGF-beta1 binding protein: a component of the large latent complex of TGF-be

A:Reference number: A35626; MUID:90275601; PMID:2350783

A:Accession: A35626

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1394 <KAN>

```

OY      232  ICGHECVNPGTYFCSCP-PGYILLDNRSQCIDNECHRNHTCNLOOTCYNLGGCFKI 290
DB      493  ICYRCINIPGSCPCSSGYSRLAPNGRCODIDECVTGIGHNSINETCNIDGAFRL 552
OY      291  DPICRCEPYLR 301
DB      553  -AECPEENYR 562

RESULT 9
T42760
fibulin, splice form D precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42760
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
submitted to the EMBL Data Library, February 1998
A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization
A:Reference number: 222267
A:Accession: T42760
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-689 <BAR>
A:Cross-references: EMBL:AF051401; PIDD: AAC28321.1
C:Genetics:
A:Note: FBLN1
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match      25.1%; Score 603; DB 2; Length 689;
Best Local Similarity 27.4%; Pred. No. 1.5e-32;
Matches 156; Conservative 52; Mismatches 155; Indels 206; Gaps 19;

OY      2  CTNGFDLDQSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPTVYRGPSNRYST 61
DB      180  CRSGFDLPADGMACVVIDECATLMDCLDSQRCILNTPGSFKCI----- 222
OY      62  PYSGPYPAAPPLSAPNPTISRLICRFGYOMD-ESNOCVDVDECATDSHOCNPQICI 120
DB      223  -----RTLSCGTGYAMDSETERCRDVECMGSHDCGPLYOCR 260
OY      121  NTEGGYTC-----SCTDGYMLLEGQCCLDIDECRYGY----- 151
DB      261  NTGGSYRCDAKKCGDGLQNPMTGECTSITCPNGYYPKNGMCMNDIDECVTGHNCGAGEC 320
OY      152  -----CGQ---LCA-----NVPSYSTCNPGE 171
DB      321  VNTPGSFRQCKGNLCAHGEYVNGATGFCEDVNECOGVCSGMECINLPETYYCKCGPGY 380
OY      172  TLNE-----DGRSCODVNECA 187
DB      381  EENDAKKRCEDVDECIKFAHGVODLSAEICNTIGSFECCKPFGOLASDGRREDVNECT 440
OY      188  TE-NPCVOTCVNTYGSFICRDPGYELEBDGVHCSMDDECSF---SEFLCOHECVNQG 242
DB      441  TGIACCEKCVNIPESYQICIDRGFALGPDGTKECEDIDECISIWAGSNDLCMGCGCINTKG 500
OY      243  TYFGSCPFGYILLDNRSQCIDNECHRNHTCNLOOTCYNLGGCFKIDIRCEEPYLR 302
DB      501  SYLCCGPGGYKIQPGRTCVDVDECA-MGECAGSDVVCVNTLGSFKC-HSIDCPNTYIHD 558
OY      303  SDNR-----C-MCPAENPGC-RDQFTILYRDMDVVSGRSV--PADI----- 340
DB      559  SLNKNQIADGVSCIVCSREDTECLGNHREVLYQRAVPSLKTIIISPIEVSRLVTHMGV 618
OY      341  -FOMQATTRYPGAYITIFQIKSGNEGREFYMRGTGPISATLWMPRIKGPRIOLDLEMIT 399
DB      619  PESVDYNDYVQGRHFRIVQERNIG-----IVQLVKPISGP-----TYETIK 660
OY      400  VNTVINFRGSSVIR-----LRIYVSQYPF 423
DB      661  VNHTKSKRTGYILANENALITELTSVSKYP 689

```

```

RESULT 10
T42990
fibulin 1, splice form C precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42990
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
submitted to the EMBL Data Library, February 1998
A:Description: Isolation of chicken and nematode fibulin-1 homologs and characterization
A:Reference number: 222267
A:Accession: T42990
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-712 <BAR>
A:Cross-references: EMBL:AF051402; PIDD: AAC28322.1
C:Genetics:
A:Gene: FBLN1
C:Superfamily: unassigned EGF-related proteins; EGF homology
C:Keywords: alternative splicing; basement membrane; extracellular matrix

Query Match      24.8%; Score 596; DB 2; Length 712;
Best Local Similarity 27.9%; Pred. No. 4.4e-32;
Matches 152; Conservative 46; Mismatches 145; Indels 202; Gaps 18;

OY      2  CTNGFDLDQSGCCLDIDECRTIPEACRGDMCVNONGYLCIPRTNPTVYRGPSNRYST 61
DB      180  CRSGFDLPADGMACVVIDECATLMDCLDSQRCILNTPGSFKCI----- 222
OY      62  PYSGPYPAAPPLSAPNPTISRLICRFGYOMD-ESNOCVDVDECATDSHOCNPQICI 120
DB      223  -----RTLSCGTGYAMDSETERCRDVECMGSHDCGPLYOCR 260
OY      121  NTEGGYTC-----SCTDGYMLLEGQCCLDIDECRYGY----- 151
DB      261  NTGGSYRCDAKKCGDGLQNPMTGECTSITCPNGYYPKNGMCMNDIDECVTGHNCGAGEC 320
OY      152  -----CGQ---LCA-----NVPSYSTCNPGE 171
DB      321  VNTPGSFRQCKGNLCAHGEYVNGATGFCEDVNECOGVCSGMECINLPETYYCKCGPGY 380
OY      172  TLNE-----DGRSCODVNECA 187
DB      381  EENDAKKRCEDVDECIKFAHGVODLSAEICNTIGSFECCKPFGOLASDGRREDVNECT 440
OY      188  TE-NPCVOTCVNTYGSFICRDPGYELEBDGVHCSMDDECSF---SEFLCOHECVNQG 242
DB      441  TGIACCEKCVNIPESYQICIDRGFALGPDGTKECEDIDECISIWAGSNDLCMGCGCINTKG 500
OY      243  TYFGSCPFGYILLDNRSQCIDNECHRNHTCNLOOTCYNLGGCFKIDIRCEEPYLR 302
DB      501  SYLCCGPGGYKIQPGRTCVDVDECA-MGECAGSDVVCVNTLGSFKC-HSIDCPNTYIHD 558
OY      303  SDNR-----C-MCPAENPGC-RDQFTILYRDMDVVSGRSV--PADI----- 337
DB      559  SLNKNRCAKROPASAGLPPE--CSKVPFLTLTYOFISL--ARAVPISHRAITLFRYSAP 613
OY      338  --ADI---FOMQATTRYPGAYITIFQIKSGNEGREFYMRGTGP--ISATLWMPRIKGP 390
DB      614  NHADDEVNFEOLKTTIIVGAPVVLPAIRAN-----FLQKGERKNSAVVTLRLSDLPQT 668
OY      391  IOLDL 395
DB      669  VKQL 673

RESULT 11
T43210
fibulin-1D precursor - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43210
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.

```


A:Residues: 30-35, 'SX', 38-40, 'SH', 43-44 <AR3>
 C:Genetics:
 A:Gene: GDB:FBLN1; FBLN
 A:Cross-references: GDB:278285; OMIM:135820
 A:Map position: 22q13.3-22q13.3
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; glycoprotein
 F:1-29/Domain: signal sequence status predicted <SIG>
 F:30-683/Product: signal sequence status predicted <SIG>
 F:180-214/Domain: EGF homology <EGF>
 F:485-523/Domain: EGF homology <EGF>
 F:98-535/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 30.6%; Score 735.5; DB 2; Length 683;
 Best Local Similarity 35.7%; Pred. No. 2,7e-41;
 Matches 155; Conservative 66; Mismatches 152; Indels 61; Gaps 16;

```

QY 1 OCTNGFDLDRSGGCLDIDEERTPEACRGDMGVNONGYLCPRTNPFYRGYSNPS 60
DB 293 QCKSGFIQD-ALGNCLIDNECLISAPCPICHTCINTGSSYTC----- 334
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOM-DESNOCVDVDECATDSHCNPQTIC 119
DB 335 -----QKNVPR-----CGRGTHNEEGTRCVDECAPPAEPCKGHRG 373
QY 120 INMEGYTCTDGYWT--LEGCLDIDEC-RY-GYCOOLCANVGSYSTCNDPFTLN 174
DB 374 VNSGSRCECKTGTYFDGISRMCVDEGCRGRLGHCNENTLGSYLSGSGFRIS 433
QY 175 EDGSCODVNECATENPVQTCVNTYGSFICRDPGELEB-DGVHCSMDMDECSF--SEF 231
DB 434 VDRGSCDINECS--SPCSQCANVGSYQCYCRGQYLSVDGVTCEIDIDECALPTEGH 492
QY 232 LCQHECVNQGTYFCSCP-PGYILLDNRSQDINECEHRNHTCNLQOTCYNLGGFCIC 290
DB 493 ICSYRCINIPGSPQSSGYSRLAPNGRNCQDIDECVTGHNCSINETCNIGCAFECCL 552
QY 291 DPICEPEYLRISDNRC-MCPA-ENPGCRDPTILYRDMVYSGRSVPADIFQWQATTR 348
DB 553 -AFECENRRSATRERLRPECHENRECSKLPRLITYHHSFPNIOAPAVFVRMGPSA 611
QY 349 YGAYVYIFQIKSGNEGREFYRQTPISATLWTRPIKGPRIOLDLDMITVNVINRG 408
DB 612 VPDSMQALITGNNEGEFFTRKRVSPHSGVALTKPVEPRDL-----LITVKMDLSNHG 666
QY 409 ----SSYIRLRIVS 419
DB 667 TVSSFVAKLFIYS 680

```

RESULT 5

A:9457
 fibulin-2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 02-Aug-2002
 C:Accession: A49457; S74095
 R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.
 J. Cell Biol. 123, 1269-1277, 1993
 A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with
 A:Reference number: A49457; MUID:94064787; PMID:8245130
 A:Accession: A49457
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:9437046; PIDN:CAAS3040.1; PID:9437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met
 A:Reference number: S74094; MUID:96439073; PMID:8841408
 A:Accession: S74095
 A:Molecule type: Protein
 A:Residues: 236-238, 'X', 240-247, 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EW', 569-589
 C:Superfamily: fibulin-2; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotr
 F:942-978/Domain: EGF homology <EGF>

Query Match 30.3%; Score 729.5; DB 2; Length 1221;
 Best Local Similarity 35.3%; Pred. No. 1.1e-40;
 Matches 145; Conservative 59; Mismatches 148; Indels 59; Gaps 11;

```

QY 1 OCTNGFDLDRSGGCLDIDEERTPEACRGDMGVNONGYLCPRTNPFYRGYSNPS 60
DB 832 RCMDFG-LDPREGCVINETSILEFCRSFCSCINTVGSYTC----- 873
QY 61 TPYSGPYPAAPPLSAPNYPITISRLICRGYOM-DESNOCVDVDECATDSHCNPQTIC 119
DB 874 -----QKNVPR-----CGRGTHNEEGTRCVDECAPPAEPCKGHRG 373
QY 120 INMEGYTCTDGYWT--LEGCLDIDECRYG--YCQOLCANVGSYSTCNDPFTLN 174
DB 913 YNIPGSTRCDCKRGFORDAFGRTCIDVNECWSVGRICQHTCENTPSSYRSCAAGFLA 972
QY 175 EDGSCODVNECATENPVQTCVNTYGSFICRDPGELEBDEGVHCSMDMDECS-FSEFIC 233
DB 973 ADGKCEDEVNECETR--CSQECANITYGSYCYGRGQLADHDHTCTDIDECAGAGATLC 1031
QY 234 QHECVNQGTYFCSCP-PGYILLDNRSQDINECEHRNHTCNLQOTCYNLGGFCIC 292
DB 1032 TTRCVNVPSSYQACPEQGTMMANGRSKDIDECALGTNCSAEATCHNIGSFRLT-R 1090
QY 293 INCEPEYLRISDNRC-MCPA-ENPGCRDPTILYRDMVYSGRSVPADIFQWQAT 346
DB 1091 FDCPPRYVAVSQTKC-----ERTTCQDITECQTSAPATITKQLNFGGLPAHIFRIGPA 1146
QY 347 TPYGAIVYIFQIKSGNEGREFYRQTPISATLWTRPIKGPRIOLDLDMITVNVINRG 408
DB 1147 PAFADDTISLITTKGNEGEFFTRKRVSPHSGVALTKPVEPRDL-----LITVKMDLSNHG 666

```

RESULT 6

A:9457
 fibulin, splice form D precursor - mouse
 M:Alternate names: basement membrane protein BM-90; calcium-binding protein BM-90
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
 C:Accession: S34968; S36441; S13814
 R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.L.
 Eur. J. Biochem. 215, 733-740, 1993
 A:Title: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-depend
 A:Reference number: S34968; MUID:93358897; PMID:8354280
 A:Accession: S34968
 A:Molecule type: mRNA
 A:Residues: 1-705 <PAN>
 R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timpl, R.; Chu, M.L.
 submitted to the EMBL Data Library, January 1993
 A:Description: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-
 A:Reference number: S36440
 A:Accession: S36441
 A:Molecule type: mRNA
 A:Residues: 1-39, 'P', 41-705 <PAN>
 A:Cross-references: EMBL:X70854; NID:9396820; PIDN:CAAS0207.1; PID:9396821
 A:Experimental source: cell-line F5 teratocarcinoma
 R:Kluge, M.; Mann, K.; Dziadek, M.; Timpl, R.
 Eur. J. Biochem. 193, 651-659, 1990
 A:Title: Characterization of a novel calcium-binding 90-kDa glycoprotein (BM-90) shar
 A:Reference number: S13814; MUID:91065369; PMID:2249686
 A:Accession: S13814
 A:Molecule type: Protein
 A:Residues: 28-31, 49, 'X', 51-53, 'XX', 110-117, 231-240, 'X', 242-243, 339-362, 'S', 364-387, 4
 C:Superfamily: fibulin-1; EGF homology
 C:Keywords: alternative splicing; basement membrane; calcium binding; extracellular m
 F:1-29/Domain: signal sequence status predicted <SIG>
 F:30-705/Product: fibulin, splice form D #status predicted <MAT>
 F:98,537,541/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

29.2%; Score 702; DB 2; Length 705;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 17:36:12 ; Search time 16.9977 Seconds
(without alignments)
2392.374 Million cell updates/sec

Title: US-09-674-379A-14

Perfect score: 2407

Sequence: 1 OCTNGFPLDRSGSCDIDIDE.....INFRSSVIRLRIVSYQYF 423

Scoring table: BLASTOM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR_73:**

1: pir1:**
2: pir2:**
3: pir3:**
4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1108.5	46.1	493	2 JC5621	epidermal growth f
2	963.5	40.0	387	2 I38449	extracellular prot
3	744.5	30.9	685	2 S78040	fibulin, splice fo
4	735.5	30.6	683	2 C36346	fibulin 1 precursor
5	729.5	30.3	1221	2 A49457	fibulin-2 precursor
6	702	29.2	705	2 S34968	fibulin, splice fo
7	701.5	29.1	1184	2 A55184	fibulin-2 precursor
8	617.5	25.1	601	2 B36346	fibulin 1 precursor
9	603	25.1	689	2 T42760	fibulin, splice fo
10	596	24.8	712	2 T42990	fibulin 1, splice
11	592.5	24.6	589	2 T43210	fibulin-1D precurs
12	574	23.8	798	2 T22793	hypothetical prote
13	539	22.4	1394	2 A35626	transforming growt
14	523.5	21.7	3002	2 A47221	fibritillin 1 precu
15	522.5	21.7	2871	2 A55567	fibritillin 1 - bovi
16	518	21.5	1712	2 A55624	masking protein pr
17	509.5	21.2	2871	2 A55624	fibritillin-1 precu
18	509.5	21.2	2907	2 A57278	fibritillin-2 precu
19	508.5	21.1	2918	2 A54105	latent transformin
20	469.5	19.5	1820	2 A55494	hypothetical prote
21	455	18.9	741	2 T46488	hypothetical prote
22	448.5	18.6	1620	2 T27283	hypothetical prote
23	438	18.2	1251	2 T52923	latent transformin
24	436	18.1	1574	2 T13954	MEGF6 protein - ra
25	392	16.3	886	2 A57172	probable hormone r
26	384.5	16.0	3507	2 T34513	hypothetical prote
27	368.5	15.3	1106	2 T18739	hypothetical prote
28	367	15.2	2471	2 A49128	cell-fate determin
29	364	15.1	810	2 T10756	Nel-homolog protei

30	361	15.0	1081	2 T31329	receptor tyrosine
31	359	14.9	1203	2 A49175	notch B protein -
32	356	14.8	2555	2 A40043	notch protein homo
33	353.5	14.7	1964	2 T09059	notch4 - mouse
34	351.5	14.6	2437	2 S42612	transmembrane prot
35	348.5	14.5	2703	1 A24420	notch protein - fr
36	340.5	14.1	673	2 A48089	growth arrest-spec
37	340	14.1	2531	2 A35844	notch protein - Af
38	335.5	13.9	2531	2 S18188	notch protein homo
39	335	13.9	674	2 T17298	growth potentialin
40	334	13.9	511	2 T17298	hypothetical prote
41	334	13.9	2531	2 T31070	notch homolog - se
42	333	13.8	1064	2 A40136	fibropellin Ia - s
43	333	13.8	1261	2 S78549	notch3 protein - h
44	325.5	13.5	2531	2 A46019	Notch-1 protein -
45	324.5	13.5	2318	2 S45306	notch 3 protein -

ALIGNMENTS

RESULT 1

JC5621.
epidermal growth factor-like protein, r16 precursor - rat

C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Oct-1997 #sequence:Revision 07-Nov-1997 #text:change 05-Nov-1999

C:Accession: JC5621

R:Ozaki, T.; Kondo, K.; Nakamura, Y.; Ichimura, S.; Nakagawara, A.; Sakiyama, S.
Biochem. Biophys. Res. Commun. 237, 245-250, 1997

A:Title: Interaction of DA41, a DAN-binding protein, with the epidermal growth factor

A:Reference number: JC5621; MUID:97415782; PMID:9268694

A:Accession: JC5621

A:Molecule type: mRNA

A:Residues: 1-493 <OZA>

A:Cross-references: DBJ:089730; NID:q2429082; PIDN:BAA22265.1; PID:di023127; PID:q24

C:Comment: This protein plays a role in the regulation of cell growth by interacting

C:Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:128-70/158-199,200-237,238-277,278-318,319-359/Region: epidermal growth factor-like

F:249/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	46.1%	Score 1108.5	DB 2	Length 493
Best Local Similarity	43.6%	Pred. No. 5e-66		
Matches 204	Conservative 68	Mismatches 149	Indels 47	Gaps 5
QY	1	OCTNGFDLDRSGSCDIDECRTIPEACRGMVQNGVLCIPRTNPNVGRPSNPS	60	
Db	28	QCTGDEWDPVHQCKDIDEDIVDACKGKCMVNHVGYCLPRTAQIIVNNQPOE	87	
QY	61	TPYS-----GPRYAAPPLSAPRYT-----	81	
Db	88	TPAABASSGAATGTAANSMTSGVITPGGFTASTAAVAGPVGRRNFVIRRPADPO	147	
QY	82	-----ISRPILCRGYQMDSESNOCVDVDECATDSHCNPTQICITMTEGGTSCDGYWL	136	
Db	148	RIPSNPSHRIGCAAGYEGSEHNVCDDICSGTINCRLDYCINLRGFTGCLPQYK	207	
QY	137	LEGGCLDIDECRY-GYCOQLCANVPGSTSCNPEFTLNEDGRSCQDVNECATENPCVQT	195	
Db	208	RGEQCVVIDECSVPYCHQGCVMTPGSFYCCNPFOIAANNYTCDVIDNECDASNOCAQO	267	
QY	196	CVNTYGSFICRDEGELEEDVHCSDMDECSFELCOHECVNOBGRYFGCSPGYILL	255	
Db	268	CYNLTGSFICOCNGYELSLRNCEDIDERTSYLCOYQVNPBPGFSCMCPQGYOV	327	
QY	256	DDNRSCDINECEHRNHTCNLQOTCYNLQGGFKCIDPRCEPYLRISDNRCMCPAENPG	315	
Db	328	-RSRTCODINECEFTNE-CREDEMKNVHGGFRCPQPNQDDPYLTSENKCVCSVTM	385	
QY	316	CRQPFILIRDMQVNGRSPVADIFQMATRRTPGATYYITQISGNGREFYMQTPPI	375	
Db	386	CRDVPQSVIYKVMIRSDRSPDIFQIATYIANTINTPRISGNENGEFYLRQTSPPV	445	

Tel: 81-92-642-5466, Fax: 81-92-642-5482)

source	location/Qualifiers
gene	1. .1536 /organism="Homo sapiens" /db_xref="taxon:9606"
CDS	1. .1536 /gene="MBP1" 97. .1428

BASE COUNT	307 a	500 c	436 g	293 t
ORIGIN				

Query Match	29.48;	Score 372.8;	DB 9;	Length 1536;
Best Local Similarity	61.38;	Pred. No. 8.5e-96;		
Matches 618; Conservative	0;	Mismatches 387;	Indels 3;	Gaps 1

[illegible]

```

Db      1018  CGCTGCGGTGGACACCAACCCGCTGCGTGGAGGCCCTGATATCCAGGTCTTGAGAACCGCTGT 1077
QY      922   ATGTGTCTCTGTGAGAACCTTGCGCTGCAGAGACAGCCCTTTACCATCTTGACCGGGAC 981
Db      1078  CTCTGCCCCGGGCTCCACCCCTCTATGTGTGAGAGACGCTTCATTCATGTGTGCACCGGTAC 1137
QY      982   ATGAGCGGTGTGTAGAGACGCTCCGTTTCCCGCTGACATCTTCGAAATTCGAAGCCACGACC 1041
Db      1138  ATGACCATTCACCTCGAGAGCGGAGCGGCCCTGACGGTTTCCAGATCCAGATCCAGCACCTCC 1197
QY      1042  CGCTACCCCTGGGGCCTATTACATTTTCCAGATTCAAATCTGGGAATGAGGGCAGAGAAATTT 1100
Db      1198  GTCTACCCCGGTGCGCTACAAATGCTTTGAGATCCGTTGTGGAAATCGCAGAGGGACCTTT 1257
QY      1102  TACATGCGGCAACAGGGGCCCATCATGATGCCACCTTGGTGATGACAGCGCCCATCAAAGG 1161
Db      1258  TACATTAGGCAAAATCAACACAGTCAGCGCCCATGTCGGTCTCGCCCGCGCGGTGACGGGC 1317
QY      1162  CCCCCGGGAATTCAGAGCTGGAGACTTGGAAATGATCTACTGTGCACACACTGTATCAACTTCAGA 1221
Db      1318  CCCCCGGGAGTACGTGTCTGGACCTGAGTAGTGTGTCAACATGAATTTCCCTATGAGCTTACCGG 1377
QY      1222  GGCAGCTTCCTGATCCGACTGCGGATATATGTGTGCGAGTACCCATCTT 1269
Db      1378  GCGCAGCTCTGTACTGAGGCTCAGCGTCTTTGTAGGGGCGCTACACCTTC 1425

```

Search completed: July 3, 2003, 15:34:28
Job time : 3310.01 secs

ACCESSION AJ132819
 VERSION AJ132819.2 GI:4456472
 KEYWORDS FBLN4 gene; fibulin-4.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1440)
 AUTHORS Glatay, R., Timml, R. and Kostka, G.
 TITLE Sequence, recombinant expression and tissue localization of two novel extracellular matrix proteins, fibulin-3 and fibulin-4
 JOURNAL Matrix Biol. 18 (5), 469-480 (1999)
 MEDLINE 20068041
 PUBLISHED 10601734
 REFERENCE 2 (bases 1 to 1440)
 AUTHORS Kostka, G.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-1999) Kosaka G., Proteinchemie, Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 82152 Martinsried, FRG
 COMMENT On Mar 21, 1999 this sequence version replaced gi:4454234.
 FEATURES
 source Location/Qualifiers
 1..1440
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="melanoma"
 1..1332
 /gene="FBLN4"
 1..1332
 /gene="FBLN4"
 /codon_start=1
 /product="fibulin-4"
 /protein_id="CA10791.2"
 /db_xref="GI:4456472"
 /db_xref="SPTREMBL:O95967"
 Translation: "MLPCTSCPLPSGLLMLALLLLGASAPQDSEBPDSYECTDYE
 WPDSDGCRDVNCLTIPKAGKEMKINHYGLICPRSAVINLIDEGEPPEPPVP
 AOHNPICPGYEPDSDSCVDVDECAQALHDCRPSODCHNLPISYOCCTPGYRKGIG
 EGVDDIECRVRYRCHRCVNPFGSCGCEFGFOLGPNRSCVDNECDMGAEQRCF
 NRYGTLCRHQGLYELHSDGSCSDIDEGSYSLCOYRCVNEPGRFCHQPOGYOL
 ARIQDIDECESGAGHOCSEAFORCVNHGSGRCVDTNRCVPEYIOVSENRCLCPASNP
 LCRBPSSIVHRTMTTSERSVPADYFOIATSYPRATYNAFOIRAGNSQDFTIRQI
 NNSAVLVARPTGPREVYLDLDMVTNLSLRASSVLKTLFVGAVTF"
 1..81
 /gene="FBLN4"
 82..1329
 /gene="FBLN4"
 /product="fibulin-4"
 BASE COUNT 291 a 467 c 397 g 285 t
 ORIGIN
 Query Match 29.4%; Score 372.8; DB 9; Length 1440;
 Best Local Similarity 61.3%; Pred. No. 8.5e-96;
 Matches 618; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

Db 565 GAGCCGGGTTTCCAGCTGGGGCCCTAACACCGCTCTGTGTGATGAACGATGTGAC 624
 QY 562 ACCGAGAACCCCTGCGCTGCAAACTGGCTTAACACCTACGGCTTTATCGCGGCT 621
 Db 625 ATGGGGGCCCCATGCGAGCAGCGCTGCTCACTCACTATGAGACCTGTGTGCTGCT 684
 QY 622 GACCCAGATATGACTTGAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
 Db 685 CACAGGGGCTATAGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 744
 QY 682 TTCTTGAGTTCTCTGCAACATGATGTGTGAACAGCCCGGACATCTTCTGCTC 741
 Db 745 TACTCAGCTACCTCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 804
 QY 742 TGCCCTCCAGGCTACATCTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 801
 Db 805 TGCCACAGGCTTACAGCTGCTG---GCCACAGGCTTGTGCAAGACATTTATGATGT 861
 QY 802 GACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
 Db 862 GAGCTGTGCTGCGACAGCAGTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
 QY 862 AATGCAATGACACCCATCCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
 Db 922 CGCTGCTGAGACACACCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
 QY 922 ATGTCCTGCTGAGACACCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
 Db 982 CTCTGCGGCTGCTGCAACCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
 QY 982 ATGAGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
 Db 1042 ATGACATACCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1101
 QY 1042 CGCTACCTGCGGCTGCTTATGATTTTCCAGTCAAAATGTGGGAATGAGGAGGAGGAGGAG 1101
 Db 1102 GTTACCTGCGGCTGCTGCAACCGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
 QY 1102 TACATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1161
 Db 1162 TACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
 QY 1162 CCCCCGGAATCCAGCTGAGTGTGAAATGATCACTGTCAACACTGTCACTCACTCA 1221
 Db 1222 CCCCCGGAATCCAGCTGAGTGTGAAATGATCACTGTCAACACTGTCACTCACTCA 1281
 QY 1222 GCGAGCTGCTGATCCGAGTGTGAAATGATCACTGTCAACACTGTCACTCACTCA 1281
 Db 1282 GCCAGCTGCTGATCCGAGTGTGAAATGATCACTGTCAACACTGTCACTCACTCA 1329
 RESULT 15
 AB030655 1536 bp mRNA linear PRI 25-MAR-2000
 LOCUS AB030655
 DEFINITION MBP1: mutant p53 binding protein 1 (MBP1), complete
 cds.
 ACCESSION AB030655
 VERSION AB030655.1 GI:7328920
 KEYWORDS MBP1: mutant p53 binding protein 1 (MBP1).
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Tanaka, S., Sugimachi, K. and Sugimachi, K.
 TITLE Human mutant p53 binding protein (MBP1)
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1536)
 AUTHORS Tanaka, S.
 TITLE Direct Submission
 JOURNAL Submitted (30-JUL-1999) Shinji Tanaka, Kyushu University, Faculty of Medicine, Department of Surgery II, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan (E-mail: shinjitesurg2.med.kyushu-u.ac.jp).

```

RESULT 13
AF137350      2304 bp      mRNA      linear      ROD 18-JAN-2000
LOCUS
DEFINITION
Rattus norvegicus embryonic vascular EGF repeat-containing protein
ACCESSION
AF137350
VERSION
AF137350.1 GI:4583508
KEYWORDS
Rattus norvegicus.
SOURCE
Rattus norvegicus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 2304)
Kowal, R.C., Richardson, J.A., Miano, J.M. and Olson, E.N.
EVEG, a novel epidermal growth factor-like repeat-containing
protein upregulated in embryonic and diseased adult vasculature
Circ. Res. 84 (10), 1166-1176 (1999)
JOURNAL
99278197
MEDLINE
10347091
PUBMED
2 (bases 1 to 2304)
Kowal, R.C., Richardson, J.A., Miano, J.M. and Olson, E.N.
Direct Submission
Submitted (25-MAR-1999) Departments of Molecular Biology and
Oncology, Internal Medicine and Pathology, UT, Southwestern Medical
Center, Dallas, 5323 Harry Hines Blvd., Dallas, TX 75235-9148
FEATURES
Source
1..2304
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
153..1499
/note="contains calcium-binding EGF repeats"
/codon_start=1
/product="embryonic vascular EGF repeat-containing protein
EVEG"
/protein_id="AAD25101.1"
/db_xref="GI:4583509"
/translating="MPSLKRLTYTTLALPHGNAOQCTNGFDDIDRQGLDID
ECRTIPACGDMACVONNGSYLCIPRTNPYRGRPNPSTYSGEYPAAPVPAS
LLESCQDIDIECRGYCOOLCANPYSCTCNGPFLINDRSCQOVNRECEPVC
QYVNTYGSFICRDPYEEDPEIDHCSMDECSFSEFLDHEVNPQPSFCCPBG
YVLEDRSCQDINECHRNHTCSPLOTCTCNGPFCIDPVEEYFLLGDNRCC
PAENTGCRDQPFILFEDMDVSGRSYPADIFQMATTRPGAYIIEIOLSGNREF
YMRGTGISLTVMTPIKPRIDQIDLEMTVNTYINFGSSVIRLRIYSQYPF"
BASE COUNT      539 a      620 c      550 g      595 t
ORIGIN
Query Match      82.3%; Score 1045; DB 10; Length 2304;
Best Local Similarity 89.0%; Pred. No. 4,7e-289;
Matches 1129; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

```

```

QY      301 GTGGATGTGAGCAGAGTGTGCAACAGATGCCACAGTCGCAACCCACCAGATCTGCATC 360
Db      528 GTGGATGTGAGCAGAGTGTGCAACAGATGCCACAGATGCCACAGTGTGCAACCCACCAGATCTGCATC 587
QY      361 AATAGTAAGGCGGGTACACCTGCTCCGACCGGAGATTTGGCTTCTGGAGGCCAG 420
Db      588 AACACGAGAGAGGATACACTGCTCTGCACTGAGGATGCTGCTTGTGGAAGGCCAG 647
QY      421 TGCTTAACATATTGATGAATGTGCTATGCTTACTGACAGCAGCTCTGTGCAATGTCT 480
Db      648 TGCTTAATATTGATGAATGTGCTATGCTTACTGACAGCAGCTCTGTGCAATGTCT 707
QY      481 GGATCCATCTCTGATACATCAACCCGTTTACCTCCATATGAGATGAGGTCTTGC 540
Db      708 GGATCCATCTCTGATACATCAACCCGTTTACCTCCATATGAGATGAGGTCTTGC 767
QY      541 CAAGATGTGACAGAGTGTGCCACCGAGAACCCCTGCTGCAAACTCTGCAACACCTAC 600
Db      768 CAAGATGTGACAGAGTGTGCAAACTGAGAACCCCTGCTGCAACACCTCTAT 827
QY      601 GGCTCTTTCATCTGCCGCTGTGACCGAGATATGACTTGAAGAATGAGGTCTTATGC 660
Db      828 GGCTCTTTCATCTGCCGCTGTGACCGAGATATGACTTGAAGAATGAGGTCTTATGC 887
QY      661 AGTATATGAGAGAGTGCAGCTTCTGAGTTCTGTCACATGATGATGTGAACAC 720
Db      888 AGTATATGAGAGAGTGCAGCTTCTGAGTTCTGTCACATGATGATGTGAACAC 947
QY      721 CCCGGACATCTCTGCTCCCTCCAGGCTACATCTCTGATGATGACCAACGAG 780
Db      948 CCCGGCTCTCTCTGCTCCCTCCAGGCTACATCTCTGATGATGATGACCAACGAG 1007
QY      781 TCCCAAGACATCAAGATATGAGACAGACAGAACACACAGCTGACAGAGAGCTG 840
Db      1008 TCCCAAGATATCAAGATATGAGACAGACAGAACACACAGCTGACAGAGCTG 1067
QY      841 TACATTTACAGAGGGGCTTCAATATGATCAGACCCATCGCTGTGAGAGGCTTATCTG 900
Db      1068 TACATTTACAGAGGGGCTTCAATATGATCAGACCCATCGCTGTGAGAGGCTTATCTG 1127
QY      901 AGGATCACTGATTAACCGCTGTATGTCTGCTGTAACCCGCTGGAGACAGACGCC 960
Db      1128 CTGATTTGGGATTAACCGCTGTATGTCTGCTGTAACCCGCTGGAGACAGACGCC 1187
QY      961 TTTCACATCTGTATACCGGAGATGAGCTGTGTGAGAGCGCTCCGTCCGCTGACATC 1020
Db      1188 TTTCACATCTGTGTTCGGAGATGATGTGTATGAGAGCGCTCCGTCCGCTGACATC 1247
QY      1021 TTCCAAATGCAAGCAGACAGCCGCTACCTGGGCTTATCATTTTCCATTTCCAAATCT 1080
Db      1248 TTCCAAATGCAAGCAGACAGCCGCTACCTGGGCTTATCATTTTCCAAATCT 1307
QY      1081 GGGATATAGGGCAGAGATTTTACATGCGGCAACCGGGCCCATCAGTGCACCTGTG 1140
Db      1308 GGGATATAGGGCAGAGATTTTACATGCGGCAACCGGGCCCATCAGTGCACCTGTG 1367
QY      1141 ATGACAGCCGCATCAAGAGGCGCCGGGAATATCAGCTGAGACTTGAATATCATCTGTC 1200
Db      1368 ATGACAGCCGCATCAAGAGGCGCCGGGAATATCAGCTGAGACTTGAATATCATCTGTC 1427
QY      1201 AACACTGTCAATCACTTACAGAGGAGCTCGGTATCGCATGCGGATATATGTGCGAG 1260
Db      1428 AACACTGTCAATCACTTACAGAGGAGCTCGGTATCGCATGCGGATATATGTGCGAG 1487
QY      1261 TACCCATTC 1269
Db      1488 TATCCGTTT 1496

```

```

RESULT 14
HSA132819      1440 bp      mRNA      linear      PRI 15-FEB-2000
LOCUS
DEFINITION
Homo sapiens mRNA for fibulin-4.

```


Db	1395	TATCCGTTTC 1403	
RESULT 11	AF112151	2478 bp mRNA linear ROD 02-AUG-1999	
LOCUS	AF112151	Mus musculus developmental arteries and neural crest EGF-like	
DEFINITION	AF112151	protein mRNA, complete cds.	
ACCESSION	AF112151	GI:5305670	
VERSION	AF112151.1	GI:5305670	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
PUBMED			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			
ORIGIN			
Query Match	83.2%	Score 1056.2; DB 10; Length 2478;	
Best Local Similarity	89.5%	Pred. No. 2.8e-292;	
Matches 1136; conservative	0;	Mismatches 133; Indels 0; Gaps 0;	
Db	1	CAGTGCAGAAATGGCTTGGACCTGATGCGCAGTGCAGACAGTGTATGATATGATGA 60	
Db	395	CAGTGCAGAAAGGGCTTTGACCTGAGACCCGACGACAGTGTATGATATGATGA 454	
QY	61	TGCGGAACATCCCGAGAGCCTGCCGAGAGACATGATGTGTTAACCAAAATGCGGG 120	
Db	455	TGCGGAGACATCCCTGAGAGCTGTGTGTGGGACATGATGTGTCAACCAAGATGCGGG 514	
QY	121	TATTTATGATTCCTCCCGGACAAACCTGTGTATGAGAGGCCCTACTCGAAACCCCTACTCG 180	
Db	515	TATTTGATGATCTCTCGAACCAACCACTGATGTGAGAGGCCCTTACTCAAAATCTACTCT 574	
QY	181	ACCCCTACTAGGTCTGACCAAGACAGTGCACCACTCTGAGCTCAAACTATGCC 240	
Db	11		


```

RESULT 9
AF093118
LOCUS Homo sapiens UP50 mRNA, complete cds.
ACCESSION AF093118
VERSION AF093118.1 GI:3676821
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Zemel, R., Sholtz, O. and Shaul, Y.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1998) Dept. of Molecular Genetics, Weizmann
Institute of Science, Rehovot 76100, Israel
FEATURES
Source
1..2019
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="urine"
1..168
169..1515
/note="50kD; similar to human SI-5 encoded by GenBank
Accession Number U03877; possible membrane and
extracellular matrix protein; urine p50 protein"
/codon_start=1
/product="p50"
/protein_id="AAC2107.1"
/db_xref="GI:3676822"
/translation="MPGIRKILVTITLALCLPSPNAOQCTNGFDLRSGOCLDID
ECRTIPKACGMVQNGVYLCNHSRTNIPYRGPSVSPYSGPPAPAPLISAP
NYPTISPLICRFQYOMDESNOCVDEQATDSHOCPICIMKGGYVSCDGYV
LLEGGCLDIDECRGYCOOLCANYPGVSCTCNGPFLINEDRCQDYNKATEPCV
QGVNTYGSFICRQDPEYELERGVHSCSDMECSFSELCHECVNPGVYFCSPG
YILLDRNSQDINECHRNHTCRLQQTGICNLQGGFCIDIRCEPEYLRISDRMC
PAENPCGRDQPTILTRMDVSVSRVPAIDFQMAATRYPGAVYIIOIKSGNGREF
YMRQGTISATLVTIRPKGREIQLDLEMITVNTVINFRSSVIRLRIVSQPF"
1516..2019
BASE COUNT 478 a 585 c 490 g 465 t 1 others
ORIGIN
3'UTR
Query Match 98.7%; Score 1253; DB 9; Length 2019;
Best Local Similarity 99.28; Pred. No. 0;
Matches 1259; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Y 1 CAGTGCAGAAATGGCTTTGACGTGATCGCAGTCAAGACAGTGTAGATTGATGAA 60
244 CAGTGCAGAAATGGCTTTGACGTGATCGCAGTCAAGACAGTGTAGATTGATGAA 303
61 TCCCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTTAAACAAATGGCGG 120
304 TCCCGAACCATCCCGAGGCGCTGCGAGAGACATGATGTGTTAAACAAATGGCGG 363
121 TATTATGATTCGCCGGACAAACCTGTGTATCGAGGGCGCTACTGAAACCCCTACTCG 180
364 TATTATGATTCGCCGGACAAACCTGTGTATCGAGGGCGCTACTGAAACCCCTACTCG 423
191 ACCCCCTACTAGTCCGTACCCAGAGCTGCCCACTCAGTCCAGTCCAACTATGCC 240
424 ACCCCCTACTAGTCCGTACCCAGAGCTGCCCACTCAGTCCAGTCCAACTATGCC 483
241 ACGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGATGAAAGCAACAATGT 300
484 ACGATCTCCAGGCGCTTATATGCGCTTTGATACAGATGATGAAAGCAACAATGT 543
301 GTGATGTGAGAGATGTGCAACAGATTTCCACAGTGAACCCCAACAGATGTGCATC 360
544 GTGATGTGAGAGATGTGCAACAGATTTCCACAGTGAACCCCAACAGATGTGCATC 603
361 AATACTGAAGCGGTACACCTGCTCTGACCGAGATATTTGGTTGGAAAGGCCAG 420
|||||

```

```

Db 604 AATATGAAGGGGGGTACACCTGCTCTGACACGAGATATGGCTTTTGGAAAGCCAG 663
QY 421 TGGTTAGACATTTGATGATGTCGCTATGTTACTGACCGAGAGCTCTGCGCAATTTCTT 480
Db 664 TGGTTAGACATTTGATGATGTCGCTATGTTACTGACCGAGAGCTCTGCGCAATTTCTT 723
QY 481 GGATCTTATCTTGTATACATGCAACCGTGTATTTACCTTCATAGAGATGAAAGTCTTGC 540
Db 724 GGATCTTATCTTGTATACATGCAACCGTGTATTTACCTTCATAGAGATGAAAGTCTTGC 783
QY 541 CAAGATGTGAACGAGTGTGCACCGAGAACCCCTGCGTGAACAACTGCGTGAACACTTAC 600
Db 784 CAAGATGTGAACGAGTGTGCACCGAGAACCCCTGCGTGAACAACTGCGTGAACACTTAC 843
QY 601 GGCTCTTTCATCTGCGCTGTGAACCCAGATATGAACTTGAAGAGATGGGCTTCATTCG 660
Db 844 GGCTCTTTCATCTGCGCTGTGAACCCAGATATGAACTTGAAGAGATGGGCTTCATTCG 903
QY 661 AGTATATGTGACAGATGACAGCTTCTCTGAGTTCTCTGCAACATGAGTGTGAACAG 720
Db 904 AGTATATGTGACAGATGACAGCTTCTCTGAGTTCTCTGCAACATGAGTGTGAACAG 963
QY 721 CCCGGACATACTTCTGCTCTGCGCTTCCAGGCTTACATCTCTGCTGATGACACCGAAGC 780
Db 964 CCCGGACATACTTCTGCTCTGCGCTTCCAGGCTTACATCTCTGCTGATGACACCGAAGC 1023
QY 781 TCCCAAGACATCAACGAATGTGACAGACAGAACACAGTGCACCAACCTGACAGACAGCTGC 840
Db 1024 TCCCAAGACATCAACGAATGTGACAGACAGAACACAGTGCACCAACCTGACAGACAGCTGC 1083
QY 841 TCAATTTTCAAGGGGGCTTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCTG 900
Db 1084 TCAATTTTCAAGGGGGCTTTCAATATGATGACCCCATCCGCTGTGAGAGACCTTATCTG 1143
QY 901 AGATATGATATTAACCGCTGTATGTCTGTCTGAGAACCCCTGCTGACAGACAGACGCC 960
Db 1144 AGATATGATATTAACCGCTGTATGTCTGTCTGAGAACCCCTGCTGACAGACAGACGCC 1203
QY 961 TTTACCATTTTACCGGGACATGACGTGTGTCTGACAGAGCTCCGTTCCGCTGACATC 1020
Db 1204 TTTACCATTTTACCGGGACATGACGTGTGTCTGACAGAGCTCCGTTCCGCTGACATC 1263
QY 1021 TTCCAAATGCAAGCCACGACCGCGCTACCTGGGGGCTATTATCAATTTTCCAGATCAATCT 1080
Db 1264 TTCCAAATGCAAGCCACGACCGCGCTACCTGGGGGCTATTATCAATTTTCCAGATCAATCT 1323
QY 1081 GGGATATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCATGTCACCCCTG 1140
Db 1324 GGGATATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGGCCCATCATGTCACCCCTG 1383
QY 1141 ATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGGAATTTGAAATGATCACTGTTC 1200
Db 1384 ATGACAGCGCCCATCAAAAGGGGCCCGGAAATCCAGCTGGAATTTGAAATGATCACTGTTC 1443
QY 1201 AACACTGTCACTCAACTTCAAGAGGAGCTCCGTATCCGACTCGGATATATGTGTCGAG 1260
Db 1444 AACACTGTCACTCAACTTCAAGAGGAGCTCCGTATCCGACTCGGATATATGTGTCGAG 1503
QY 1261 TACCATTTTC 1269
Db 1504 TACCATTTTC 1512

```

```

RESULT 10
BC006636
LOCUS Mus musculus, fibulin 5, clone MGC:5656 IMAGE:3482574, mRNA,
DEFINITION complete cds.
ACCESSION BC006636
VERSION BC006636.1 GI:13879321
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus

```

QY 961 TTATACATCTTGTACCGGAGATGAGAGTGTGTGAGAGAGCTCCGTTCCCGCTGACATC 1020
|||||
Db 1485 TTATACATCTTGTACCGGAGATGAGAGTGTGTGAGAGAGCTCCGTTCCCGCTGACATC 1544
QY 1021 TTCCAAATGCAAGCCAGACGACCGCTTACCTGGGGCTTATTAATTTTCCAGATCAAAATCT 1080
|||||
Db 1545 TTCCAAATGCAAGCCAGACGACCGCTTACCTGGGGCTTATTAATTTTCCAGATCAAAATCT 1604
QY 1081 GGGAAATGAGGGAGAGAAATTTTACATGCGGGCAAAAGGGCCCATCAGTGGCCACCTGGTG 1140
|||||
Db 1605 GGGAAATGAGGGAGAGAAATTTTACATGCGGGCAAAAGGGCCCATCAGTGGCCACCTGGTG 1664
QY 1141 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATGATCACTGTC 1200
|||||
Db 1665 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATGATCACTGTC 1724
QY 1201 AACACTGTATCACTTACAGAGGAGCTCCGTGATCCGACTGGCGAATATATGTGCGAG 1260
|||||
Db 1725 AACACTGTATCACTTACAGAGGAGCTCCGTGATCCGACTGGCGAATATATGTGCGAG 1784
QY 1261 TACCCATTC 1269
|||||
Db 1785 TACCCATTC 1793

RESULT 8
AX403659 2609 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 14 from Patent W00077037.
AX403659
VERSION AX403659.1 GI:21437116
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS
1 Ashkenazi, A., Baker, K., Botstein, D., Desnoyers, L., Eaton, D.L.,
Ferrara, N., Fong, S., Gao, W.Q., Gerber, H., Gerritsen, M.E.,
Goddard, A., Godowski, P., Gurney, A., Kijavich, I.J., Mather, J.,
Napier, M., Pan, J., Paoni, N., Roy, M., Tumas, D., Watanabe, C.,
Williams, P.M., Wood, W.I. and Zhang, Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
the same
Patent: WO 0077037-A 14 21-DEC-2000;
Genentech Inc. (US)

FEATURES
Source Location/Qualifiers
1.2609
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 622 a 718 c 622 g 647 t

ORIGIN
Query Match 99.9%; Score 1267.4; DB 6; Length 2609;
Best Local Similarity 99.9%; Pired. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 ACATCTCCAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGAACCAATGT 300
|||||
Db 765 ACATCTCCAGGCTCTTATATGCGGCTTGGATACAGATGATGAAGAACCAATGT 824
QY 301 GTGGATGTGAGCAGATGTGCAACAGATTCACACAGTGCACACCCAGATCTGCATC 360
|||||
Db 825 GTGGATGTGAGCAGATGTGCAACAGATTCACACAGTGCACACCCAGATCTGCATC 884
QY 361 AATACGGAAGGCGGTACACTGCTCTGACACGACAGATATTTGGCTTCTGGAAGCCAG 420
|||||
Db 885 AATACGGAAGGCGGTACACTGCTCTGACACGACAGATATTTGGCTTCTGGAAGCCAG 944
QY 421 TCGTTAGACATTTATGATGTGCTATGATGTTAGTACGACAGCTCTGTGCAATGTTCT 480
|||||
Db 945 TCGTTAGACATTTATGATGTGCTATGATGTTAGTACGACAGCTCTGTGCAATGTTCT 1004
QY 481 GGATCTATTTCTTATCATGCAACCCGTGTTTACCTCAATGAGATGAAAGTCTTGC 540
|||||
Db 1005 GGATCTATTTCTTATCATGCAACCCGTGTTTACCTCAATGAGATGAAAGTCTTGC 1064
QY 541 CAAGATGTGACAGATGTGCAACCCGTGTTTACCTCAATGAGATGAAAGTCTTGC 600
|||||
Db 1065 CAAGATGTGACAGATGTGCAACCCGTGTTTACCTCAATGAGATGAAAGTCTTGC 1124
QY 601 GGCTCTTCTATCTGCGCTGTGACCCAGATATGAACTTGAAGATGAGGCTTCTTGC 660
|||||
Db 1125 GGCTCTTCTATCTGCGCTGTGACCCAGATATGAACTTGAAGATGAGGCTTCTTGC 1184
QY 661 AGTATATGAGCAGATGTGCTGTGAGTCTCTGCAACATGATGTGTGAACAG 720
|||||
Db 1185 AGTATATGAGCAGATGTGCTGTGAGTCTCTGCAACATGATGTGTGAACAG 1244
QY 721 CCCGGACATACCTTCTGCTCTGCTGCGCTGACAGGCTACATCTGAGATGACAAACGAGC 780
|||||
Db 1245 CCCGGACATACCTTCTGCTCTGCTGCGCTGACAGGCTACATCTGAGATGACAAACGAGC 1304
QY 781 TGCAGACATCAACGATGTGAGCAGAGAACACAGTGCACACCTGACAGAGCTGC 840
|||||
Db 1305 TGCAGACATCAACGATGTGAGCAGAGAACACAGTGCACACCTGACAGAGCTGC 1364
QY 841 TACAAATTTACAAAGGGGCTTCAAAATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 900
|||||
Db 1365 TACAAATTTACAAAGGGGCTTCAAAATGATGATGATGACCCCATCCGCTGTGAGAGGCTTATCTG 1424
QY 901 AGGATGATGATACCGCTGTATGTCTGCTGAGAACCCCTGGCTGACAGAGCAGCC 960
|||||
Db 1425 AGGATGATGATACCGCTGTATGTCTGCTGAGAACCCCTGGCTGACAGAGCAGCC 1484
QY 961 TTATACATCTTGTACCGGAGATGAGAGTGTGTGAGAGAGCTCCGTTCCCGCTGACATC 1020
|||||
Db 1485 TTATACATCTTGTACCGGAGATGAGAGTGTGTGAGAGAGCTCCGTTCCCGCTGACATC 1544
QY 1021 TTCCAAATGCAAGCCAGACGACCGCTTACCTGGGGCTTATTAATTTTCCAGATCAAAATCT 1080
|||||
Db 1545 TTCCAAATGCAAGCCAGACGACCGCTTACCTGGGGCTTATTAATTTTCCAGATCAAAATCT 1604
QY 1081 GGGAAATGAGGGAGAGAAATTTTACATGCGGGCAAAAGGGCCCATCAGTGGCCACCTGGTG 1140
|||||
Db 1605 GGGAAATGAGGGAGAGAAATTTTACATGCGGGCAAAAGGGCCCATCAGTGGCCACCTGGTG 1664
QY 1141 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATGATCACTGTC 1200
|||||
Db 1665 ATGACAGCGCCCATCAAGAGGCGCCGGGAAATCCAGCTGAGCTTGAATATGATCACTGTC 1724
QY 1201 AACACTGTATCACTTACAGAGGAGCTCCGTGATCCGACTGGCGAATATATGTGCGAG 1260
|||||
Db 1725 AACACTGTATCACTTACAGAGGAGCTCCGTGATCCGACTGGCGAATATATGTGCGAG 1784
QY 1261 TACCCATTC 1269
|||||
Db 1785 TACCCATTC 1793

QY 541 CAAGATGTGAAGAGTGTGTCACCGAGAACCCCTGCTGCAAAACCTGGCTCAACACTTAC 600
|||||
Db 997 CAAGATGTGAAGAGTGTGTCACCGAGAACCCCTGCTGCAAAACCTGGCTCAACACTTAC 1056
QY 601 GGCCTTTTCATCTGCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCTTCAATTGC 660
|||||
Db 1057 GGCCTTTTCATCTGCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCTTCAATTGC 1116
QY 661 AGTATATGAGCAGTGTGACCTTCTGAGTTTCCCTGTCGCAACATGAGTGTGGAACCG 720
|||||
Db 1117 AGTATATGAGCAGTGTGACCTTCTGAGTTTCCCTGTCGCAACATGAGTGTGGAACCG 1176
QY 721 CCGGCGCAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
Db 1177 CCGGCGCAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
QY 781 TGGCAACATCAACGAATGTGAGCAGACGAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 840
|||||
Db 1237 TGGCAACATCAACGAATGTGAGCAGACGAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1296
QY 841 TACAAATTTACAAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATG 900
|||||
Db 1297 TACAAATTTACAAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATG 1356
QY 901 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
|||||
Db 1357 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1416
QY 961 TTTTACCATCTTGTACCGGGAGATGAGCTGTGTAGAGACGCTGCTGCTGCTGCTGCTGCTGCTG 1020
|||||
Db 1417 TTTTACCATCTTGTACCGGGAGATGAGCTGTGTAGAGACGCTGCTGCTGCTGCTGCTGCTGCTG 1476
QY 1021 TTCCAAATGCAAGGCGACAGACCCGCTACCCCTGGGGCTTTTACATTTTCCAGATCAAAATCT 1080
|||||
Db 1477 TTCCAAATGCAAGGCGACAGACCCGCTACCCCTGGGGCTTTTACATTTTCCAGATCAAAATCT 1536
QY 1081 GGGATGAGGGGCGAGAGATTTTACATGCGGCAAAAGGGGCGCCATCAGTGCACCCCTGTGTG 1140
|||||
Db 1537 GGGATGAGGGGCGAGAGATTTTACATGCGGCAAAAGGGGCGCCATCAGTGCACCCCTGTGTG 1596
QY 1141 ATGACAGGCGCCCATCAAAAGGGGCGCCGGAATCCAGCTGTGAGCTTTGGAATGATCACTGTG 1200
|||||
Db 1597 ATGACAGGCGCCCATCAAAAGGGGCGCCGGAATCCAGCTGTGAGCTTTGGAATGATCACTGTG 1656
QY 1201 AACAGCTCATCACTTCAGAGGAGCTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCG 1260
|||||
Db 1657 AACAGCTCATCACTTCAGAGGAGCTCCGATCCGATCCGATCCGATCCGATCCGATCCGATCCG 1716
QY 1261 TACCCATTC 1269
|||||
Db 1717 TACCCATTC 1725

RESULT 7
AX376340 2609 bp DNA linear PAT 01-MAR-2002
LOCUS AX376340
DEFINITION Sequence 407 from Patent WO0168848.
ACCESSION AX376340
VERSION
KEYWORDS
SOURCE AX376340.1 GI:19170561
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Baker, K. P., Chen, J., Desnoyers, L., Goddard, A., Godowski, P. J.,
Gurney, A. L., Pan, J., Smith, V., Watanabe, C. K., Wood, W. I. and
Zhang, Z.
TITLE Secreted and transmembrane polypeptides and nucleic acids encoding
the same
JOURNAL Patent: WO 0168848-A 407 20-SEP-2001;
Genentech, Inc. (US)
FEATURES Location/Qualifiers

source 1. .2609
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 622 a 718 c 622 g 647 t
ORIGIN
Query Match 99.9%; Score 1267.4; DB 6; Length 2609;
Best local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CAGTGCACGAATGGCTTTGACCTGATGCGCAGTCCAGACAGTGTATGATGATGAA 60
|||||
Db 525 CAGTGCACGAATGGCTTTGACCTGATGCGCAGTCCAGACAGTGTATGATGATGAA 584
QY 61 TGGCGAACAATCCCGAGGGCTGCGGAGAGACTGATGTGTAAACAAATGGGGGG 120
|||||
Db 585 TGGCGAACAATCCCGAGGGCTGCGGAGAGACTGATGTGTAAACAAATGGGGGG 644
QY 121 TATTTATGATTCCTCCGGACAAACCTGTGTATGAGAGGGCTTACTGAAACCTTACTG 180
|||||
Db 645 TATTTATGATTCCTCCGGACAAACCTGTGTATGAGAGGGCTTACTGAAACCTTACTG 704
QY 181 ACCCCCTACTCAGGTCCTGACCCAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
|||||
Db 705 ACCCCCTACTCAGGTCCTGACCCAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
QY 241 AGGATCAGGCTGCTTATATGCGCGCTTGTGATGACAGATGATGATGATGATGATGATGATGATG 300
|||||
Db 765 AGGATCAGGCTGCTTATATGCGCGCTTGTGATGACAGATGATGATGATGATGATGATGATGATG 824
QY 301 GTGATGTGAGCAGTGTGCAACAGATTTCCACAGTGTGCAACCCACCCAGATGTGATC 360
|||||
Db 825 GTGATGTGAGCAGTGTGCAACAGATTTCCACAGTGTGCAACCCACCCAGATGTGATC 884
QY 361 AATACGAAAGGGGGGTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
|||||
Db 885 AATACGAAAGGGGGGTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
QY 421 TGCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
|||||
Db 945 TGCTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004
QY 481 GGATCCTATTCCTGTATGATGCAACCCCTGTGTTTACCTCAATGATGATGATGATGATGATG 540
|||||
Db 1005 GGATCCTATTCCTGTATGATGCAACCCCTGTGTTTACCTCAATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAAGAGTGTGTCACCGAGAACCCCTGCTGCAAAACCTGGCTCAACACTTAC 600
|||||
Db 1065 CAAGATGTGAAGAGTGTGTCACCGAGAACCCCTGCTGCAAAACCTGGCTCAACACTTAC 1124
QY 601 GGCCTTTTCATCTGCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCTTCAATTGC 660
|||||
Db 1125 GGCCTTTTCATCTGCGCTGTGACCCAGATATGAACTTTGAGGAAGATGGCTTCAATTGC 1184
QY 661 AGTATATGAGCAGTGTGACCTTCTGAGTTTCCCTGTCGCAACATGAGTGTGGAACCG 720
|||||
Db 1185 AGTATATGAGCAGTGTGACCTTCTGAGTTTCCCTGTCGCAACATGAGTGTGGAACCG 1244
QY 721 CCGGCGCAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
Db 1245 CCGGCGCAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
QY 781 TGGCAACATCAACGAATGTGAGCAGACGAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 840
|||||
Db 1305 TGGCAACATCAACGAATGTGAGCAGACGAAACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 1364
QY 841 TACAATTTCAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 900
|||||
Db 1365 TACAATTTCAAGGGGGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1424
QY 901 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
|||||
Db 1425 AGGATCAGTATTAACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484

```

QY 61 TGGCAACATCCCGAGGCTGCGAGGAGACATGATGTGTTAACCAAAATGGCGG 120
DB 517 TGGCAACATCCCGAGGCTGCGAGGAGACATGATGTGTTAACCAAAATGGCGG 576
QY 121 TATTATGATCCCGAGAACCCCTGTGTATGAGGGCCCTACTCGAACCCCTACTG 180
DB 577 TATTATGATCCCGAGAACCCCTGTGTATGAGGGCCCTACTCGAACCCCTACTG 636
QY 181 ACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCCTCTCAGTCCAAATATCC 240
DB 637 ACCCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCCTCTCAGTCCAAATATCC 696
QY 241 ACATCTCCAGGCTCTTATATGCGGCTTGTGATACGAGATGAAAGCAACCAATG 300
DB 697 ACATCTCCAGGCTCTTATATGCGGCTTGTGATACGAGATGAAAGCAACCAATG 756
QY 301 GTGGATGTGAGCAGTGTGCACAGATTCACCACAGTGCACACCCACCAGATCTGCATC 360
DB 757 GTGGATGTGAGCAGTGTGCACAGATTCACCACAGTGCACACCCACCAGATCTGCATC 816
QY 361 AATACGTAAGGGGGGTACACCTGCTCCTGCACCGAGATATTGGCTTCTGGAAGCCAG 420
DB 817 AATACGTAAGGGGGGTACACCTGCTCCTGCACCGAGATATTGGCTTCTGGAAGCCAG 876
QY 421 TGTGTAGACATGTGATGATGTGCTATGTACTGTCCAGCAGCTCTGTGCGAATGTCT 480
DB 877 TGTGTAGACATGTGATGATGTGCTATGTACTGTCCAGCAGCTCTGTGCGAATGTCT 936
QY 481 GGATCTTATCTTGTATACATCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 540
DB 937 GGATCTTATCTTGTATACATCAACCCCTGTGTTTACCTCAATGAGATGGAAGTCTTGC 996
QY 541 CAAGATGTGACAGTGTGCACCGAGAACCCCTGCGTGCACAACTGGGTCAACACTTAC 600
DB 997 CAAGATGTGACAGTGTGCACCGAGAACCCCTGCGTGCACAACTGGGTCAACACTTAC 1056
QY 601 GGCTCTTTCATCTGCCGCTGTGACCCAGATATCACTTGAAGAATGGCGTCTATTCG 660
DB 1057 GGCTCTTTCATCTGCCGCTGTGACCCAGATATCACTTGAAGAATGGCGTCTATTCG 1116
QY 661 AGTGAATGTGACAGTGTGCACCTGTGTAGTCTCTGCCAATCATGAGTGTGTAACCG 720
DB 1117 AGTGAATGTGACAGTGTGCACCTGTGTAGTCTCTGCCAATCATGAGTGTGTAACCG 1176
QY 721 CCGGACATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 1177 CCGGACATACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
QY 781 TGGCAAGACATCAAGATGTGACAGAGAGAACCAACAGTGCACACTGACAGACAGTGC 840
DB 1237 TGGCAAGACATCAAGATGTGACAGAGAGAACCAACAGTGCACACTGACAGACAGTGC 1296
QY 841 TACAAATTTACAAGGGGCTTCAAAATGCATGCAACCCCATCCGCTGTGAGAGCCCTATCTG 900
DB 1297 TACAAATTTACAAGGGGCTTCAAAATGCATGCAACCCCATCCGCTGTGAGAGCCCTATCTG 1356
QY 901 AGGATAGTATTAACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 1357 AGGATAGTATTAACCGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1416
QY 961 TTTACCATCTTGTACCGGAGATGAGAGTGTGTACAGACGCTCCGTTCCCGCTGACATC 1020
DB 1417 TTTACCATCTTGTACCGGAGATGAGAGTGTGTACAGACGCTCCGTTCCCGCTGACATC 1476
QY 1021 TTTCAAAATGTAAGCAGACCCGCTACCCCTGAGGCTTATTAATTTTCCAGATCAATCT 1080
DB 1477 TTTCAAAATGTAAGCAGACCCGCTACCCCTGAGGCTTATTAATTTTCCAGATCAATCT 1536
QY 1081 GGAATAGAGGAGAGAAATTTTACATGCGGCAAGGGGCCCATAGTGGCCACCTGGTG 1140
DB 1537 GGAATAGAGGAGAGAAATTTTACATGCGGCAAGGGGCCCATAGTGGCCACCTGGTG 1596

```

```

QY 1141 ATGACAGCCCCCATCAAAAGGGCCCCGGGAAATCCAGTGGACTTGGAAATGATCACTGC 1200
DB 1597 ATGACAGCCCCCATCAAAAGGGCCCCGGGAAATCCAGTGGACTTGGAAATGATCACTGC 1656
QY 1201 AACACTGTCATCAACTTCAGAGGAGAGTCCGTGATCCGATCCGATATATGTGTGCAG 1260
DB 1657 AACACTGTCATCAACTTCAGAGGAGAGTCCGTGATCCGATCCGATATATGTGTGCAG 1716
QY 1261 TACCATTCT 1269
DB 1717 TACCATTCT 1725

RESULT 6
AR173204
LOCUS AR173204 2550 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 2 from patent US 6303765.
ACCESSION AR173204
VERSION AR173204.1 GI:17912695
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2550)
AUTHORS Bandman, O., Corley, N.C. and Guegler, K.J.
TITLE Human extracellular matrix proteins
JOURNAL Patent: US 6303765-A 2 16-OCT-2001;
FEATURES
Source 1. 2550
/organism="Unknown"

BASE COUNT 623 a 690 c 596 g 641 t
ORIGIN

Query Match 100.0%; Score 1269; DB 6; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


FEATURES passed the following selection criteria: matched mRNA gi: 5453649.
source Location/Qualifiers

1..2367
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22412 IMAGE:4693953"
/tissue_type="Lung"
/clone_id="NTH.MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
184..1530
/product="Unknown (protein for MGC:22412)"
/protein_id="AAH2280.1"
/db_xref="gi:18490145"
/translation="MPSIKRLTYITLALCLPSPQMAOCTNFDLDRSGCCLDID
ECRTIPAGKDMKCVNONGYLCIPRNPYRGPSPYSPYSGPAPAPLSAP
NYPISPLICREYOMDESNOVDVCATDSHCNFTQICINTEGGYSCITDGY
LLEGOCADIDECRYCOOLCANYPGSGCNPFILNEDRSCODVNECATENPCV
OTCVNTYGSFICRCDPELEEDGHSMDMECSFSEFLCOHECNOPGTFCSCPC
YLLDNRSCODINECHRNHTCNLOOTCVNLOGSPFCIDPICEPEPLRISDNCKC
PAENPGCRDOPFTLYRDMVVSRSVPADIFOMATTRIRGATYIFQIKSGNEREF
YMRQTPISATLVMTPIKPREIQLDLEMTVNTVINFRSSVIRLTIYVSQYPF"

BASE COUNT 588 a 650 c 547 g 582 t
ORIGIN

Query Match 100.0%; Score 1269; DB 9; Length 2367;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAGACAGTGTATGATATGATGAA 60
DB 259 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAGACAGTGTATGATATGATGAA 318
QY 61 TGCCGAACCATCCCGAGGCGCTGCGGAGACATGATGTGTGTTAACCAAAATGGCGGG 120
DB 319 TGCCGAACCATCCCGAGGCGCTGCGGAGACATGATGTGTGTTAACCAAAATGGCGGG 378
QY 121 TATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 180
DB 379 TATTTATGCAATCCCGGACAAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTCG 438
QY 181 ACCCCCTACTGAGTGGCTGATCCAGCAGCTGCGCCACCACTCTGACCTCCAACTATCC 240
DB 439 ACCCCCTACTGAGTGGCTGATCCAGCAGCTGCGCCACCACTCTGACCTCCAACTATCC 498
QY 241 AGCATCTCCAGGCTCTTATATGCCCTTGTGATACAGATGATGTAAGAACAACATGT 300
DB 499 AGCATCTCCAGGCTCTTATATGCCCTTGTGATACAGATGATGTAAGAACAACATGT 558
QY 301 GTGGATGTGGAGAGTGTGACACAGATTTCCACCACTGACACCCACCAAGATCTGCATC 360
DB 559 GTGGATGTGGAGAGTGTGACACAGATTTCCACCACTGACACCCACCAAGATCTGCATC 618
QY 361 AATACTGAAGGCGGATGACACCTGCTCTGACACGAGGATATGGCTTGGAAAGCCAG 420
DB 619 AATACTGAAGGCGGATGACACCTGCTCTGACACGAGGATATGGCTTGGAAAGCCAG 678
QY 421 TGTCTAGACATGATGAATGTGCTATGTTACTGACAGAGCTGTGCGAATGTTCTT 480
DB 679 TGTCTAGACATGATGAATGTGCTATGTTACTGACAGAGCTGTGCGAATGTTCTT 738
QY 481 GGATCTCTTCTTGTGATGACAAACCTGTTTACCTCAATGAGATGGAAGTCTTTC 540
DB 739 GGATCTCTTCTTGTGATGACAAACCTGTTTACCTCAATGAGATGGAAGTCTTTC 798
QY 541 CAAGATGTGAAGGATGTGCGACCGAGACCCCTGCTGCAAAACCTGCGGCAACACCTTAC 600
DB 799 CAAGATGTGAAGGATGTGCGACCGAGACCCCTGCTGCAAAACCTGCGGCAACACCTTAC 858
QY 601 GGTCTTTTCATCTGCGGCTGTGACCCAGGATATGAACCTGAGGAAGATGGCGTTTATTC 660
DB 859 GGTCTTTTCATCTGCGGCTGTGACCCAGGATATGAACCTGAGGAAGATGGCGTTTATTC 918

```

```

QY 661 AGTATATGAGCAGATGACAGCTTCTCTGAGTCTCTGCGCAACATGATGATGTGAACAG 720
DB 919 AGTATATGAGCAGATGACAGCTTCTCTGAGTCTCTGCGCAACATGATGATGTGAACAG 978
QY 721 CCCGGCAGATATCTTCTGCTCTGCGGCTGCGGCTGAGTACATCTGATGATGACACGAGC 780
DB 979 CCCGGCAGATATCTTCTGCTCTGCGGCTGCGGCTGAGTACATCTGATGATGACACGAGC 1038
QY 781 TGCCAGACATCAACGAGATGTGACACAGAGACACACAGGCAACCTGCGACACAGCTGC 840
DB 1039 TGCCAGACATCAACGAGATGTGACACAGAGACACACAGGCAACCTGCGACACAGCTGC 1098
QY 841 TACAAATTTACAAAGGGGCTTCAATATGATGACACCCATCCGCTGTGAGAGACCTTATCTG 900
DB 1099 TACAAATTTACAAAGGGGCTTCAATATGATGACACCCATCCGCTGTGAGAGACCTTATCTG 1158
QY 901 AGGATGAGATATTAACCGCTGTATGTCTGCTGAGAACCTGCTGCGACAGACACCC 960
DB 1159 AGGATGAGATATTAACCGCTGTATGTCTGCTGAGAACCTGCTGCGACAGACACCC 1218
QY 961 TTTACCATCTTGTACCGGACATGACAGTGTGTGACAGAGCTCCGCTGAGACATC 1020
DB 1219 TTTACCATCTTGTACCGGACATGACAGTGTGTGACAGAGCTCCGCTGAGACATC 1278
QY 1021 TTCCAAATGCAAGCCACGACCGCTTACCTGCGGCTTATTAATTTTCCAGATCAATCT 1080
DB 1279 TTCCAAATGCAAGCCACGACCGCTTACCTGCGGCTTATTAATTTTCCAGATCAATCT 1338
QY 1081 GGAATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCATTCATGATGACACCTGTG 1140
DB 1339 GGAATGAGGCGCAGAGATTTTACATGCGGCAAAAGGCGCCATTCATGATGACACCTGTG 1398
QY 1141 ATGACAGCGCCCATCAAAAGGCGCCGGAATTCACAGTGTGAAATGATGATGATGTC 1200
DB 1399 ATGACAGCGCCCATCAAAAGGCGCCGGAATTCACAGTGTGAAATGATGATGATGTC 1458
QY 1201 AACATGTCATCACTTCAGAGGAGCTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1260
DB 1459 AACATGTCATCACTTCAGAGGAGCTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1518
QY 1261 TACCAGATTC 1269
DB 1519 TACCAGATTC 1527

```

RESULT 5

AR036548 2550 bp DNA linear PAT 29-SEP-1999
LOCUS AR036548
DEFINITION Sequence 2 from patent US 5,872,234.
ACCESSION AR036548
VERSION AR036548.1 GI:5953216

SOURCE

ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 2550)
AUTHORS Bandman O., Corley N.C. and Guegler K.J.
TITLE Human extracellular matrix proteins
JOURNAL Patent: US 5,872,234-A 2 16-FEB-1999;
FEATURES Location/Qualifiers

source 1..2550
/organism="unknown"

BASE COUNT 623 a 690 c 596 g 641 t
ORIGIN

Query Match 100.0%; Score 1269; DB 6; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAGACAGTGTATGATATGATGAA 60
DB 457 CAGTGCACGAATGGCTTTGACCTGGATGCGCAGTCAGACAGTGTATGATATGATGAA 516

```


crest cells, balloon injured vessels, atherosclerotic lesions" /codon_start=1 /product="developmental arteries and neural crest EGF-like protein" /protein_id="AAD41768.1" /db_xref="GI:5305673" /translation="MGCIRLLPTLALICSPNMAOCTNGFDIDRSGCCLDIECRTIPACRGDMACVNONGSYLCIPETNTPYRKPINPSTPISGTPPAAPLSAIPNPLISPLICRFGYQMDNESNOCVDVDECATDSHQCPNQLCIWTEGGYTCSDGFWLLEGGCLDIDBERGRYGOQLCANAPGVSCTCNGEFTLINBEGRCQDVNDECATNPPCQQTVCNNGYSFICRCDPSGELEEDHSCSDMECSFSELCQHECVNODPGATFCSCPGAYILDDNRSCODINECHRNHTCNLQOTCYNLGGFPCIDIRCEEPYLRISDNKRCQYIAENPGCRGDDPFTLYRDMDNVVSGRSYPAIDFQMDATRYDGAIVYIOLSGNKGREELYMNOTGCISSLTIVTPRIKRPKEPREDQLDEMLTAVTVINFRSSVIRLRITVYSQTPP"

Query Match	100.0%	Score 1269;	DB 9;	length 2328;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 1269;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0

Y	1	CAGGCGACGATGGCTTTGACCTGGATCGCCAGCCAGACAGTGTATTGATTTGATGAA	60
D	244	CAGTGCACGATGGCTTTGACCTGGATCGCCAGCCAGACAGTGTATTGATTTGATGAA	303
QY	61	TGCGGAACCAATCCCGGAGGCGCTGCCGAGGAGACATGATGTGTATTAAACAAATAGCGGG	120
D	304	TGCGGAACCAATCCCGGAGGCGCTGCCGAGGAGACATGATGTGTATTAAACAAATAGCGGG	363
QY	121	TATTATGCAATTCGCGGACAAACCTGTGTATGAGGGGCCCTACTGMAACCCCTACTCG	180
D	364	TATTATGCAATTCGCGGACAAACCTGTGTATGAGGGGCCCTACTGMAACCCCTACTCG	423
QY	181	ACCCCCACAGGTGCGTACCCAGCAGCTGCCACACACTCTAGCTCCAACTATATCC	240
D	424	ACCCCCACAGGTGCGTACCCAGCAGCTGCCACACACTCTAGCTCCAACTATATCC	483
QY	241	ACGATCTCAGGCGCTTATATGCGCGTTTGATACAGATGATGAAAGCAACCAATGT	300
D	484	ACGATCTCAGGCGCTTATATGCGCGTTTGATACAGATGATGAAAGCAACCAATGT	543
QY	301	GTGATGTGACGACGATGTGCACACAGATTTCCACAGTGCACCCACCAGATCTGATC	360
D	544	GTGATGTGACGACGATGTGCACACAGATTTCCACAGTGCACCCACCAGATCTGATC	603
QY	361	AATCTGAAGCGGGTACACCTGCTCCGACGACGAGATTTTGGCTTCGGAAGGCCAG	420
D	604	AATCTGAAGCGGGTACACCTGCTCCGACGACGAGATTTTGGCTTCGGAAGGCCAG	663
QY	421	TGCTTAGACATTTGTAATGTGCGTATGTGTTACTGACGACGCTGTGCGCAATGTTCCT	480
D	664	TGCTTAGACATTTGTAATGTGCGTATGTGTTACTGACGACGCTGTGCGCAATGTTCCT	723
QY	481	GGATCTATTTCTGTACATGCACCCCTGGTGTAACTCAATGAGAGTGAAGTCTTGC	540
D	724	GGATCTATTTCTGTACATGCACCCCTGGTGTAACTCAATGAGAGTGAAGTCTTGC	783
QY	541	CAAGATGTGAACGAGTGTGCCACCGGAAACCCCTGCGTGAAACCTGCGTAAACACTAC	600
D	784	CAAGATGTGAACGAGTGTGCCACCGGAAACCCCTGCGTGAAACCTGCGTAAACACTAC	843
QY	601	GCGCTTTCAATTCGCGTGACCCAGAGATATACCTTGAGGAAGATGGGCTCATTTGC	660
D	844	GCGCTTTCAATTCGCGTGACCCAGAGATATACCTTGAGGAAGATGGGCTCATTTGC	903
QY	661	AGTGAATGACGAGTGCACCTTCTGAGTTCCTCGCAACATGATGTGTGAACACG	720
D	904	AGTGAATGACGAGTGCACCTTCTGAGTTCCTCGCAACATGATGTGTGAACACG	963
QY	721	CCGGGACATATCTTCTGCTGCTGCCCTCCAGGCTACATCTGCTGATGACACCGAAGC	780
D	964	CCGGGACATATCTTCTGCTGCTGCCCTCCAGGCTACATCTGCTGATGACACCGAAGC	1023

OY	781	GGCCAGACATCAACGATGTGAGACACAGGAACCAACGATCACTCGGACACACGTCG	840
Db	1024	TGCCAAGACATCAACGAATGTGACACAGGAACCAACGATCACTCGGACACACGTCG	1083
OY	841	TACAAATTACAAAGGGGGCTTCAAATGCAATGCATGCACCCATCCGCTGTGAGGAGCCCTTATCTG	900
Db	1084	TACAATTACAAAGGGGGCTTCAAATGCAATGCATGCACCCATCCGCTGTGAGGAGCCCTTATCTG	1143
OY	901	AGGATCAAGTATAAACCGCTGTATGTGTCTCTGTGAAACCTGGCTGCAGAGACCAGCC	960
Db	1144	AGGATCAAGTATAAACCGCTGTATGTGTCTCTGTGAAACCTGGCTGCAGAGACCAGCC	1203
OY	961	TTTACCATCTGTACCGGGAGCATGTGAGCTGTGTGCAGGACGCTCCGTTCCCGCTGACATC	1020
Db	1204	TTTACCATCTGTGTACCGGGAGCATGTGAGCTGTGTGTGCAGGACGCTCCGTTCCCGCTGACATC	1263
OY	1021	TTTCCAAATGCAAGCCAGCCAGCCGCTACCCCTGGGGCCTATTACATTTTCCAGATCAAACT	1080
Db	1264	TTTCCAAATGCAAGCCAGCCAGCCGCTACCCCTGGGGCCTATTACATTTTCCAGATCAAACT	1323
OY	1081	GGGAATGAGGGCCACAGAAATTTTACATGTGGGGAACGGGCCCATCAGTCCACCTGGTG	1140
Db	1324	GGGAATGAGGGCCACAGAAATTTTACATGTGGGGAACGGGCCCATCAGTCCACCTGGTG	1383
OY	1141	ATGACACGCCCCATCAAAAGGGGCCCGGGAATTCAGCTGTGACTTGGAATGATCACTGTC	1200
Db	1384	ATGACACGCCCCATCAAAAGGGGCCCGGGAATTCAGCTGTGACTTGGAATGATCACTGTC	1443
OY	1201	AACACTGTCAATCAACTCAGAGGCAAGCTCGGTGATCCGATCGCGGATATATGTGTCCGAG	1260
Db	1444	AACACTGTCAATCAACTCAGAGGCAAGCTCGGTGATCCGATCGCGGATATATGTGTCCGAG	1503
OY	1261	TACCCATTTC 1269	
Db	1504	TACCCATTTC 1512	
RESULT 4			
LOCUS	BC022280	2367 bp	mRNA linear PRI 04-FEB-2002
DEFINITION	Homo sapiens, clone MGC:22412 IMAGE:4693955, mRNA, complete cds.		
ACCESSION	BC022280		
VERSION	BC022280.1	GI:18490144	
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2367)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-FEB-2002) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		

RESULT 4					
BC022280					
LOCUS	BC022280	2367 bp	mRNA	linear	PRI 04-FEB-2002
DEFINITION	Homo sapiens, clone MGC:22412 IMAGE:4693955,				
ACCESSION	BC022280				complete cds.
VERSION	BC022280.1				
KEYWORDS	GI:18490144				
SOURCE	MGC.				
ORGANISM	Homo sapiens.				
	Homo sapiens				
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 2367)				
AUTHORS	Strausberg,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-FEB-2002) National Institutes of Health, Mammalian				
	Gene Collection (MGC), Cancer Genomics Office, National Cancer				
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,				
	USA				
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				
COMMENT	Contact: MGC help desk				

Email: cgabps-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdeparill@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 36 Row: n Column: 23
This clone was selected for full length sequencing because it

```

/protein_id="CA838568.1"
/db_xref="GI:4490530"
/translation="MSEIKRLIVTITIALCLPSPGNMAOQCTNEDIDLRSGCLDID
ECHTIPACGDMVQNGVGLCTIPRTNPYRGPSPYSPYSGPPAPAPLSAP
NPTPLICRFQYOMDESNOCDVDECATDSHOQNTQICINTEGGYCTSGTGY
LLEGOCIDIDECRGYCOOLCANVPSYSCNPFITNEGRSGODVNEATEPCV
QTCVNTGSGFICRDPXELBEDGVHCSMDDECSFSEFICQHEVNOGYTFCSPG
YILLDNRSCODINECBRHHTCNLOOTCYNLOGSPFCIDIRCEBPLRISDNKCM
PANKRCRDPFTILYRDMVVSRSVPADIFOMATTPGAVIFQIKSGNEREF
YMRQPTISATLVMTFRKIPRETQIDLEMTVNTVFNFRGSVIRLRIVYSQYF"
sig_peptide
1..66
/gene="FIBL-5"
/mat_peptide
67..1344
/gene="FIBL-5"
/product="fblin-5"
BASE COUNT      507 a      540 c      475 g      496 t      1 others
ORIGIN
Query Match      100.0%; Score 1269; DB 9; Length 2019;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CAGTGACGAAATGGCTTTGACCTGATCCGACAGTGTATGATATGATGAA 60
|||||
76 CAGTGACGAAATGGCTTTGACCTGATCCGACAGTGTATGATATGATGAA 135
|||||
61 TCCCGAACCATCCCGAGGCGCTCCGAGAGACATGATGTGTATACCAAAATGGGGG 120
|||||
136 TCCCGAACCATCCCGAGGCGCTCCGAGAGACATGATGTGTATACCAAAATGGGGG 195
|||||
121 TATTTATGATTCGCCGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTCG 180
|||||
196 TATTTATGATTCGCCGACAAACCTGTGTATGAGGGCCCTACTGAAACCCCTACTCG 255
|||||
181 ACCCCCTACTAGTCCGTACCCAGACAGCGCCCGACCACTCTCAGTCCCAACTATCCC 240
|||||
256 ACCCCCTACTAGTCCGTACCCAGACAGCGCCCGACCACTCTCAGTCCCAACTATCCC 315
|||||
241 ACGATCTCAGGCGCTTATATGCCCTTTGATACAGATGATGAAACCAACATGT 300
|||||
316 ACGATCTCAGGCGCTTATATGCCCTTTGATACAGATGATGAAACCAACATGT 375
|||||
301 GTGATGTGACAGATGTGCAACAGATTTCCACAGTGCACACCCACCGATCTGCATC 360
|||||
376 GTGATGTGACAGATGTGCAACAGATTTCCACAGTGCACACCCACCGATCTGCATC 435
|||||
361 AATACTGAAGGGGTACACCTGCTCTGACACGAGGATATGCTTCGGAAGGCAG 420
|||||
436 AATACTGAAGGGGTACACCTGCTCTGACACGAGGATATGCTTCGGAAGGCAG 495
|||||
421 TGCTTAGACATTTGATGATGTGCTATGCTTACCTGCACAGCTCTGTGGAATGTTCCT 480
|||||
496 TGCTTAGACATTTGATGATGTGCTATGCTTACCTGCACAGCTCTGTGGAATGTTCCT 555
|||||
481 GGATCTCTTCTGTGACATGCAACCTGGTTTACCTCAATGAGATGGAAGCTTTCG 540
|||||
556 GGATCTCTTCTGTGACATGCAACCTGGTTTACCTCAATGAGATGGAAGCTTTCG 615
|||||
541 CAAAGATGTAAGAGTGTGCAACGAGAACCCCTGCTCAAAACCTGCTCAACACTTC 600
|||||
616 CAAAGATGTAAGAGTGTGCAACGAGAACCCCTGCTCAAAACCTGCTCAACACTTC 675
|||||
601 GGCTCTTCTGATCTGCGCTGTGACACGAGATATGAATCTGAGAGAGTGGCTTATTCG 660
|||||
676 GGCTCTTCTGATCTGCGCTGTGACACGAGATATGAATCTGAGAGAGTGGCTTATTCG 735
|||||
661 AGTGAATGAGAGAGTGTGCTGCTGAGTTCCTGCTGCAACATGAGTGTGAACGAG 720
|||||
736 AGTGAATGAGAGAGTGTGCTGCTGAGTTCCTGCTGCAACATGAGTGTGAACGAG 795
|||||
721 CCGGACACATATCTGCTCTGCTCCAGGCTACATCTCTGATGACAAACGAGAC 780
|||||
796 CCGGACACATATCTGCTCTGCTCCAGGCTACATCTCTGATGACAAACGAGAC 855

```

```

QY 781 TGCCAGACATCAACGAATGTGACAGCAGGACACAGCTGCAACCTGCAAGCAGCTGC 840
|||
DB 856 TGCCAGACATCAACGAATGTGACAGCAGGACACAGCTGCAACCTGCAAGCAGCTGC 915
|||
QY 841 TACAAATTCAGAGGGGGCTTCAAAATGCAATGACACCCATCCGCTGTGAGAGACCTTATCTG 900
|||
DB 916 TACAAATTCAGAGGGGGCTTCAAAATGCAATGACACCCATCCGCTGTGAGAGACCTTATCTG 975
|||
QY 901 AGGATGAGATTAACGGCTGTATGTCTCCGCTGAGAGAACCTGGCTGCAGAGACACCC 960
|||
DB 976 AGGATGAGATTAACGGCTGTATGTCTCCGCTGAGAGAACCTGGCTGCAGAGACACCC 1035
|||
QY 961 TTTACCATCTTGTACCGGGACATGAGAGCTGTGTGTCAGAGAGCTCCGCTTCCCGTGACATC 1020
|||
DB 1036 TTTACCATCTTGTACCGGGACATGAGAGCTGTGTGTCAGAGAGCTCCGCTTCCCGTGACATC 1095
|||
QY 1021 TTCCAAATTCAGAGCCACGACCCGCTTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1080
|||
DB 1096 TTCCAAATTCAGAGCCACGACCCGCTTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1155
|||
QY 1081 GGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCTCAGTGCACACCTGGTG 1140
|||
DB 1156 GGAATGAGGGGACAGAAATTTTACATGCGGCAAAAGGGGCCCTCAGTGCACACCTGGTG 1215
|||
QY 1141 ATGACAGCCCGCATCAAGAGGGCCCGGGAATTCAGCTGACTTGGAAATGATGACTGTC 1200
|||
DB 1216 ATGACAGCCCGCATCAAGAGGGCCCGGGAATTCAGCTGACTTGGAAATGATGACTGTC 1275
|||
QY 1201 AACACTGCATCAACTTTCAGAGGAGCTCGTATCCGAGTGGGATATATGTGTGCGAG 1260
|||
DB 1276 AACACTGCATCAACTTTCAGAGGAGCTCGTATCCGAGTGGGATATATGTGTGCGAG 1335
|||
QY 1261 TACCACTTC 1269
|||
DB 1336 TACCACTTC 1344
|||
RESULT 3
AF112152 2328 bp mRNA linear PRI 02-AUG-1999
LOCUS Homo sapiens developmental arteries and neural crest EGF-like
DEFINITION protein mRNA, complete cds.
ACCESSION AF112152
VERSION AF112152.1 GI:5305672
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1. (bases 1 to 2328)
AUTHORS Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Tanikawa, M.,
Matsunori, A., Sasayama, S., Chien, K. R. and Honjo, T.
TITLE DANCE, a novel secreted RGD protein expressed in developing
adrenoleiomyoblasts and leiomyoma-injured arteries
JOURNAL St. Biol. Chem. 274 (492), 22476-22488 (1999)
MEDLINE 99357749
PUBMED 10428823
REFERENCE 2. (bases 1 to 2328)
AUTHORS Nakamura, T., Yabe, D., Tashiro, K. and Honjo, T.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1998) Medicine, University of California San
Diego, 9500 Gilman Drive, La Jolla, CA 92093-0613, USA
FEATURES
Source
1..2328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="14"
/map="14q32.1"
169..1515
/feature=DANCE; contains 6 cbef domains and 1 RGD motif;
secreted protein; expressed in developmental aorta, neural

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:43:26 ; Search time 3309.01 Seconds

(Without alignments)
11160.894 Million cell updates/sec

File: US-09-674-379A-15

Sequence: 1 cagtgacgacgatgcttga.....atgtgcgcagctaccatc 1269

Scoring table: IDENTITY_NIC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank
1: gb|0000000000
2: gb|0000000000
3: gb|0000000000
4: gb|0000000000
5: gb|0000000000
6: gb|0000000000
7: gb|0000000000
8: gb|0000000000
9: gb|0000000000
10: gb|0000000000
11: gb|0000000000
12: gb|0000000000
13: gb|0000000000
14: gb|0000000000
15: gb|0000000000
16: gb|0000000000
17: gb|0000000000
18: gb|0000000000
19: gb|0000000000
20: gb|0000000000
21: gb|0000000000
22: gb|0000000000
23: gb|0000000000
24: gb|0000000000
25: gb|0000000000
26: gb|0000000000
27: gb|0000000000
28: gb|0000000000
29: gb|0000000000
30: gb|0000000000
31: gb|0000000000
32: gb|0000000000
33: gb|0000000000
34: gb|0000000000
35: gb|0000000000
36: gb|0000000000
37: gb|0000000000
38: gb|0000000000
39: gb|0000000000
40: gb|0000000000
41: gb|0000000000

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1269	100.0	1720	6	A84086 Sequence 1
2	1269	100.0	2328	9	AJ133490 Homo sapi
3	1269	100.0	2367	9	AF112152 Homo sapi
4	1269	100.0	2550	6	BC022280 Homo sapi
5	1269	100.0	2550	6	AR036548 Homo sapi
6	1269	100.0	2550	6	AR173204 Homo sapi
7	1267.4	99.9	2609	6	AX376340 Sequence
8	1267.4	99.9	2609	6	AX403659 Sequence
9	1253	98.7	2019	9	AF093118 Homo sapi
10	1056.2	83.2	2230	10	BC006636 Mus muscu
11	1056.2	83.2	2478	10	AF112151 Mus muscu
12	1046.6	82.5	2499	10	AF112153 Mus muscu
13	1045	82.3	2304	10	AF137350 Mus muscu
14	372.8	29.4	1440	9	HSA133819 Homo sapi
15	372.8	29.4	1536	9	AB030655 Homo sapi
16	371.2	29.3	1480	6	AX023967 Sequence
17	371.2	29.3	1561	9	BC010456 Homo sapi
18	371.2	29.3	1707	9	AK000980 Homo sapi
19	371.2	29.3	1875	6	AX201326 Homo sapi
20	371.2	29.3	1875	6	AX464184 Homo sapi
21	371.2	29.3	2018	6	AR095382 Homo sapi
22	369.6	29.1	1757	9	AF109121 Homo sapi
23	364.8	28.7	1825	9	AF124486 Homo sapi
24	363.8	28.7	1705	9	AF093119 Homo sapi
25	350.4	27.6	1778	10	AF046870 Homo sapi
26	347.2	27.4	1513	10	AF104223 Homo sapi
27	347.2	27.4	1513	6	AX023961 Sequence
28	345.6	27.2	1794	10	AF109122 Homo sapi
29	345.6	27.2	1724	10	BC012265 Homo sapi
30	337.6	26.6	1358	6	AX023976 Sequence
31	312	24.6	1021	6	AX023954 Sequence
32	295.8	23.3	2613	9	AK098186 Homo sapi
33	268	21.1	1122	6	AX023965 Sequence
34	260.6	20.5	171901	9	AC007057 Homo sapi
35	260.6	20.5	172674	9	CNS00008 Homo sapi
36	260.6	20.5	175210	9	CNS07EG5 Homo sapi
37	250.6	19.7	1677	10	BC031184 Homo sapi
38	249	19.6	1911	10	DB9730 Homo sapi
39	241	19.0	2025	9	BC014410 Homo sapi
40	241	19.0	2512	6	AX277602 Sequence
41	241	19.0	2512	6	AX331916 Sequence
42	241	19.0	2512	6	AX332783 Sequence
43	241	19.0	2512	6	AX333462 Sequence
44	241	19.0	2512	6	AX334103 Sequence
45	241	19.0	2512	6	AX335899 Sequence

ALIGNMENTS

RESULT 1
LOCUS A84086 1720 bp DNA
DEFINITION Sequence 1 from Patent WO9846746.
ACCESSION A84086
VERSION A84086.1 GI:6733224
KEYWORDS
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1720)
AUTHORS Li, H. and Olsen, H. S.
TITLE EXTRACELLULAR/EPIDERMAL GROWTH FACTOR LIKE PROTEIN
JOURNAL Patent: WO 9846746-A 1 22-OCT-1998;
HUMAN GENOME SCIENCES INC (US); LI HAODONG (US)

```
QY 1021 TTCGAATGCAAGCCAGACCCGCTACCTGGGGCTATTACATTTTCAGATCAAAATCT 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1096 TTCAGATGCAAGCAACAACCCGATACCTGGTGCTATTACATTTTCAGATCAAAATCT 1155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1081 GGGAAATGAGGGCAGAGAAATTTTACATGCGGCAAAAGGGCCCCCATCAGTCCACCCCTGGTG 1140
    || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1156 GGCACAGAGGGTCTGAGAGTTCTATATGCGGCAAAAGGGCCCTATCAGTCCACCCCTGGTG 1215
    || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1141 ATGACAGGCCCCCATCAAGGGCCCCGGGAATCCAGCTGACCTGGAAATGATCACTGTC 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1216 ATGACAGGCCCCCATCAAGGGCCCCGGGAATCCAGCTGACCTGGAGATGATCACTGTC 1275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1201 AACACTGTATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCAG 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1276 AACACTGTATCAACTTTCAGAGGAGCTCCGTGATCCGACTGCGGATATATGTGCGCAG 1335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1261 TACCCATTC 1269
    || || ||||
Db 1336 TATCCGTTT 1344
    || || ||||
```

Search completed: July 3, 2003, 11:59:31
Job time : 283.77 secs

Db 1022 TCTGGGAATGAGGGGAGAGATTTTTCATGCGGCAAAACGGGCCCATCATCTGCACCCCTG 1081
 QY 1138 GTGATGACACGCCCCCATCAAGAGGCCCGGGAATTCAGCTGACTTGGAAATGATCACT 1197
 Db 1082 GTGATGACACGCCCCCATCAAGAGGCCCGGGAATTCAGCTGACTTGGAAATGATCACT 1141
 QY 1198 GTCAACACGTGATCAATCTTCAAGAGGACCTCGTATCTCCGACTGCGGATATATGTGTCG 1257
 Db 1142 GTCAACACGTGATCAATCTTCAAGAGGACCTCGTATCTCCGACTGCGGATATATGTGTCG 1201
 QY 1258 CAGTACCCATTC 1269
 Db 1202 CAGTACCCATTC 1213

RESULT 13
 AA239385
 AA239385 standard: DNA: 1269 BP.

AA239385;

22-FEB-2000 (first entry)

Smooth muscle proliferation modulating mature protein coding sequence.

Cell proliferation; vascular; smooth muscle cell; arteriosclerosis; PTCA; endothelial thickening; percutaneous transluminal coronary angioplasty; myoma; hematopoietic cell-regulation; cytokine; tissue generation; taxis; actin activity; chemotaxis; blood coagulation; thrombotic; tumor; mouse; metastasis; nutrient; ss.

Mus musculus.

MO9955863-A1.

04-NOV-1999.

28-APR-1999: 99MO-JP02283.

28-APR-1998: 98JP-0119731.

(ONOX) ONO PHARM CO LTD.

Honjo T, Tashiro K, Nakamura T;

MP1: 2000-038646/03.

P-PSDB: AA156751.

Novel mouse polypeptides for treatment, e.g. arteriosclerosis and myoma

Claim 4: Page 53-54; 70pp; Japanese.

The invention provides mouse polypeptides for treatment of diseases due to abnormal proliferation of smooth muscle. The polypeptides can be produced by standard recombinant methodology. The polypeptides can be used according to their inhibition of the proliferation of vascular smooth muscle cells, particularly in treating arteriosclerosis or retransluminar coronary angioplasty (PTCA), or myoma, hematopoietic cell-regulatory activity, cytokine activity, tissue generation/repair, coagulation/thrombotic activity, receptor/ligand activity, cadherin/actin activity, chemotaxis activity, tumor inhibition, and as nutrient. The present sequence represents the coding sequence of the mature protein cell proliferation.

Sequence 1269 BP; 304 A; 350 C; 314 G; 301 T; 0 other.

Query Match 83.2%; Score 1056.2; DB 21: Length 1269;
 Best Local Similarity 89.5%; Pred. No. 0;

Matches 1136; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 1 CAGTGCAGAAATGGCTTTGACCTGGATCGCAGTCAAGACAGTGTATGATATGAGAA 60
 Db 1 CAGTGCAGAAATGGCTTTGACCTGGATCGCAGTCAAGACAGTGTATGATATGAGAA 60
 QY 61 TGGCGAAGCAATCCCGGAGGCTCGCGAGAGACATGATGTGTAAACAAATGGCGGG 120
 Db 61 TGGCGAAGCAATCCCGGAGGCTCGCGAGAGACATGATGTGTAAACAAATGGCGGG 120
 QY 121 TATTTAGCATTTCCCGGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTCG 180
 Db 121 TATTTAGCATTTCCCGGAGAAACCTGTGTATGAGAGGCTTACTGAAACCTTACTCG 180
 QY 181 ACCCCCTACAGTCCGATCCAGACAGAGCTGCCACACTCTCAGTCCAAATATCC 240
 Db 181 ACATCTCTACAGGCCCATACCCAGAGAGGCGCCACAGTACACCTTCCAACTACCC 240
 QY 241 ACAGATCCAGGCTCTTATATGCGCTTGTGATACCAAGATGATGAAAGCAACCAATGT 300
 Db 241 ACAGATCCAGGCTCTTGTCTGCGCTTGTGATACCAAGATGATGAAAGCAACCAATGT 300
 QY 301 GTGATGTGAGAGAGTGTGCAACAGATTCACCAAGTGCACCCACAGATCTGATAC 360
 Db 301 GTGATGTGAGAGAGTGTGCAACAGATTCACCAAGTGCACCCACAGATCTGATAC 360
 QY 361 AATATGAAAGGCGGTACACCTGCTCTGACAGAGAGATATGCTTCTGGAAGCCAG 420
 Db 361 AATATGAAAGGCGGTACACCTGCTCTGACAGAGAGATATGCTTCTGGAAGCCAG 420
 QY 421 TGCCTTAGACATGATGAATGTGCTATGCTGCTGACAGACCTGTGCAATGTCT 480
 Db 421 TGCCTTAGATGATGAATGTGCTATGCTGCTGACAGACCTGTGCAATGTCT 480
 QY 481 GGATCCATTTCTTGTACATGGAACCCGTGTTTACCCTCAATGAGATGGAAGTCTTGC 540
 Db 481 GGATCCATTTCTTGTACATGGAACCCGTGTTTACCCTCAATGAGATGGAAGTCTTGC 540
 QY 541 CAAGATGTGAAGAGTGTGCCAGGAGAACCCCTGCTGCAAACTGGGTCAACACCTAC 600
 Db 541 CAAGATGTGAAGAGTGTGCCAGGAGAACCCCTGCTGCAAACTGGGTCAACACCTAC 600
 QY 601 GGCCTTTTCATCTGCGCTGCTGACCCAGGATATGACATTTGAGAAAGATGCGCTTATTC 660
 Db 601 GGCCTTTTCATCTGCGCTGCTGACCCAGGATATGACATTTGAGAAAGATGCGCTTATTC 660
 QY 661 AGTATATGAGAGAGTGTGACAGTCTTGTGAGTCTCTGCAACATGATGTGTACACAG 720
 Db 661 AGTATATGAGAGAGTGTGACAGTCTTGTGAGTCTCTGCAACATGATGTGTACACAG 720
 QY 721 CCGGCGACATCTTCTGCTGCTGCTGCTGACCCAGGCTTACATCTGTGATGATACCAAGC 780
 Db 721 CCGGCGCATCTTCTGCTGCTGCTGCTGACCCAGGCTTACATCTGTGATGATACCAAGC 780
 QY 781 TGGCAAGACATCAACGATGTGACACAGAGACACAGAGTGCAGACAGTGC 840
 Db 781 TGGCAGAGATTCATGATGTGACACAGAGACACAGAGTGCAGACAGTGC 840
 QY 841 TACAAATTTACAGGCGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
 Db 841 TACAAATTTACAGGCGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 901 AGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 901 AGGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 961 TTTTACCATTTTACCGGAGATGAGAGTGTGTGAGAGAGTGTGTGAGAGAGTGTGTGAGAG 1020
 Db 961 TTTTACCATTTTACCGGAGATGAGAGTGTGTGAGAGAGTGTGTGAGAGAGTGTGTGAGAG 1020
 QY 1021 TTCCAAATGAG 1080
 Db 1021 TTCCAAATGAG 1080

22-MAY-2000; 2000MO-US14042.
PR 15-JUN-1999; 99US-0139695.
PR 20-JUL-1999; 99US-0145070.
PR 26-AUG-1999; 99US-0145698.
PR 17-AUG-1999; 99US-0149369.
PR 01-SEP-1999; 99MO-US20111.
PR 08-SEP-1999; 99MO-US20594.
PR 15-SEP-1999; 99MO-US21090.
PR 15-SEP-1999; 99MO-US21547.
PR 30-NOV-1999; 99MO-US28313.
PR 01-DEC-1999; 99MO-US28301.
PR 02-DEC-1999; 99MO-US28565.
PR 07-DEC-1999; 99US-0169495.
PR 05-JAN-2000; 2000MO-US00219.
PR 18-FEB-2000; 2000MO-US04341.
PR 18-FEB-2000; 2000MO-US04342.
PR 22-FEB-2000; 2000MO-US04414.
PR 01-MAR-2000; 2000MO-US05601.
PR 02-MAR-2000; 2000MO-US05841.
PR 20-MAR-2000; 2000MO-US07377.
PR 30-MAR-2000; 2000MO-US08439.
PR 15-MAY-2000; 2000MO-US13358.
PR 17-MAY-2000; 2000MO-US13705.

(GETH) GENENTECH INC.

PI Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gao W, Gerber H, Gertlisen ME, Goddard A;
PI Godowski PJ, Guney AL, Kijavlin IU, Mather JP, Napier MA, Pan J;
PI Picon NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;
PI Wood WI, Zhang Z;

WPI; 2001-050091/06.
P-PSDB; AAB31183.

PT Isolated nucleic acid molecule encoding a PRO polypeptide which is a
transmembrane polypeptide is useful for gene therapy and identification
of related polypeptides -

Claim 2; Fig 9; 244pp; English.

PS The present sequence encodes a human secreted and transmembrane
XX polypeptide. The specification describes human polypeptides, designated
CC PRO16, PRO44, PRO183, PRO185, PRO210, PRO215, PRO217, PRO242, PRO288,
CC PRO365, PRO353, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,
CC PRO348, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
CC PRO337, PRO341, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells
CC can be modulated with agents that bind to these polypeptides, resulting
CC in the death of the cells. The polynucleotides encoding these
CC polypeptides are useful in the recombinant production of the
CC polypeptides, as a hybridisation probe to screen libraries to isolate
CC homologous sequences, or to map the gene. They may also be used for
CC analysing genetic disorders, and to produce transgenic animals which are
CC useful for the development and screening of therapeutically useful
CC reagents. The polynucleotides can also be used in gene therapy e.g. to
CC replace a defective gene.

XX Sequence 2609 BP; 622 A; 718 C; 622 G; 647 T; 0 other;

Query Match 99.98; Score 1267.4; DB 22; Length 2609;
Best Local Similarity 99.98; Pred. No. 0;
Matches 1288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAGTCACGATGCTTTGACCTGATGCGGACGACGAGAGCTTATGATGATGAA 60
DB 525 CAGTCACGATGCTTTGACCTGATGCGGACGACGAGAGCTTATGATGATGAA 584
QY 61 TGGCAGACCATCCCGAGGCGTCCGAGAGACATGATGTTAAACCAAAATGGCGG 120
DB 585 TGGCAGACCATCCCGAGGCGTCCGAGAGACATGATGTTAAACCAAAATGGCGG 644

QY 121 TATTATGATCCCGGACAAACCCCTGTATGAGGGCCCTACTCGAACCCCTACTCG 180
DB 645 TATTATGATCCCGGACAAACCCCTGTATGAGGGCCCTACTCGAACCCCTACTCG 704
QY 181 ACCCCCTACTGAGTCCGTACCCAGACAGCTGCCCCAGCAGCTGCTGCAATATGCC 240
DB 705 ACCCCCTACTGAGTCCGTACCCAGACAGCTGCCCCAGCAGCTGCTGCAATATGCC 764
QY 241 ACGATCTCAGGCGCTTATATGCGCGCTTGGATACGAGATGATGAAGCAACCAATGT 300
DB 765 ACGATCTCAGGCGCTTATATGCGCGCTTGGATACGAGATGATGAAGCAACCAATGT 824
QY 301 GTGATGTGAGAGGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGCATC 360
DB 825 GTGATGTGAGAGGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGCATC 884
QY 361 AATACGTAAGGCGGCTACCTGCTGCGACCGAGATATGCTGCTGGAAGCGCAG 420
DB 885 AATACGTAAGGCGGCTACCTGCTGCGACCGAGATATGCTGCTGGAAGCGCAG 944
QY 421 TGCTTAGACATTTGATGATGCTGCTATGTTACTGCCAGAGCTGTGGAAATGCTCT 480
DB 945 TGCTTAGACATTTGATGATGCTGCTATGTTACTGCCAGAGCTGTGGAAATGCTCT 1004
QY 481 GGATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 540
DB 1005 GGATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAAGAGGTGTCACAGACGACGACGACGACGACGACGACGACGACGACG 600
DB 1065 CAAGATGTGAAGAGGTGTCACAGACGACGACGACGACGACGACGACGACGACGACG 1124
QY 601 GGCCTCTTCATCTGCGCTGTGACCCAGATATGATGATGATGATGATGATGATG 660
DB 1125 GGCCTCTTCATCTGCGCTGTGACCCAGATATGATGATGATGATGATGATGATG 1184
QY 661 AGTATATGAGAGGAGGACCTTCTGATGATGATGATGATGATGATGATGATGATG 720
DB 1185 AGTATATGAGAGGAGGACCTTCTGATGATGATGATGATGATGATGATGATGATG 1244
QY 721 CCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 1245 CCGGACATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
QY 781 TGGCAGACATCAGCAATGTGAGCAGCAGACCAACAGTGCACCTGACAGACGTCG 840
DB 1305 TGGCAGACATCAGCAATGTGAGCAGCAGACCAACAGTGCACCTGACAGACGTCG 1364
QY 841 TACAATTTCAAGGGGGCTTCMAATGATGATGATGATGATGATGATGATGATGATG 900
DB 1365 TACAATTTCAAGGGGGCTTCMAATGATGATGATGATGATGATGATGATGATGATG 1424
QY 901 AGGATCAGTATTAACCGCTGTATGCTGCTGAGAACCCCTGCTGAGAGACCAAGCCC 960
DB 1425 AGGATCAGTATTAACCGCTGTATGCTGCTGAGAACCCCTGCTGAGAGACCAAGCCC 1484
QY 961 TTTACCATCTTACCCGAGACATGACGAGCTGTGACGAGCTCTCCGCTGACATC 1020
DB 1485 TTTACCATCTTACCCGAGACATGACGAGCTGTGACGAGCTCTCCGCTGACATC 1544
QY 1021 TTCCAAATGACAGCCAGACCGCTACCTGCGGCTATTAATTTTCCAGATCAATCT 1080
DB 1545 TTCCAAATGACAGCCAGACCGCTACCTGCGGCTATTAATTTTCCAGATCAATCT 1604
QY 1081 GGGATGAGGCGAGAAATTTTACATGCGGCAAGCGGCGCCATCAGTGCACACCTGTG 1140
DB 1605 GGGATGAGGCGAGAAATTTTACATGCGGCAAGCGGCGCCATCAGTGCACACCTGTG 1664
QY 1141 ATGACAGCCCGCATCAAGAGGCGCCCGGGAATTCAGCTGTGGAATATGATCTGTC 1200
DB 1665 ATGACAGCCCGCATCAAGAGGCGCCCGGGAATTCAGCTGTGGAATATGATCTGTC 1724
QY 1201 AACACTGTATCACTTCAAGAGGCGCTCGATCCGAGCTCGGATATATGTGTGCGAG 1260

to screen for modulators of the compounds -
 Claim 2; Fig 407; 774pp; English.

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

Sequence 2609 BP; 622 A; 718 C; 622 G; 647 T; 0 other:

Query Match 99.98; Score 1267.4; DB 22; Length 2609;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CAGTGCAGCAATGGCTTTGACCTGGATCGCCAGTCCAGGACAGTGTATGATTTGATGAA 60
DB 525 CAGTGCAGCAATGGCTTTGACCTGGATCGCCAGTCCAGGACAGTGTATGATTTGATGAA 584
OY 61 TCCCGAACCATCCCGAGGCCCTGCCAGAGACATGATGTGTAAACCAAAATGGCGGG 120
DB 585 TCCCGAACCATCCCGAGGCCCTGCCAGAGACATGATGTGTAAACCAAAATGGCGGG 644
OY 121 TATTATGCAATTCCTCCGGAGAAAACCTGTGTATGAGGGCCCTACTGCAACCCCTACTG 180
DB 645 TATTATGCAATTCCTCCGGAGAAAACCTGTGTATGAGGGCCCTACTGCAACCCCTACTG 704
OY 181 ACCCCCTACTGAGGTCCGTACCCAGCAGCTGCCCCACACTCTGCTGCAAACTATGCC 240
DB 705 ACCCCCTACTGAGGTCCGTACCCAGCAGCTGCCCCACACTCTGCTGCAAACTATGCC 764
OY 241 ACGATCTCCAGGCGCTTTATATGCGGCTTTGGATACCAAGATGATGAAGCAACCAATGT 300
DB 765 ACGATCTCCAGGCGCTTTATATGCGGCTTTGGATACCAAGATGATGAAGCAACCAATGT 824
OY 301 GTGGATGTGACAGTGTGCAACAGATTTCCACAGTGCACCCCAACCAATGTGCTATC 360
DB 825 GTGGATGTGACAGTGTGCAACAGATTTCCACAGTGCACCCCAACCAATGTGCTATC 884
OY 361 AATACGTAAGGCGGCTACACCTGCTCGACAGCAGGATTTGGCTTTGGAAGGCCAG 420
DB 885 AATACGTAAGGCGGCTACACCTGCTCGACAGCAGGATTTGGCTTTGGAAGGCCAG 944
OY 421 TCGTTAGACATGATGAATGCGGTATGTTACTGCCAGCAGCTCTGTGCGAATGTCTCT 480
DB 945 TCGTTAGACATGATGAATGCGGTATGTTACTGCCAGCAGCTCTGTGCGAATGTCTCT 1004
OY 481 GGATCCTATTTCTGTACATGCAACCCCTGTTTACCTTCATGAGATGGAAGGCTTGC 540
DB 1005 GGATCCTATTTCTGTACATGCAACCCCTGTTTACCTTCATGAGATGGAAGGCTTGC 1064
OY 541 CAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTCAAACTGCGTCAACACATAC 600
DB 1065 CAAGATGTGAACAGTGTGCGACCGAGAACCCCTGCGTCAAACTGCGTCAACACATAC 1124
OY 601 GCGCTCTTCAATCTGCGCTGTGACCCAGATATGAACCTTGAAGAGATGGCGTTCAATGC 660
DB 1125 GCGCTCTTCAATCTGCGCTGTGACCCAGATATGAACCTTGAAGAGATGGCGTTCAATGC 1184
OY 661 AGGATATGAGAGAGTGCACCTTCTGAGTCTCTGCGCAACATGATGTGTGAACAG 720
DB 1185 AGGATATGAGAGAGTGCACCTTCTGAGTCTCTGCGCAACATGATGTGTGAACAG 1244

```

```

DB 1185 AGGATATGAGAGAGTGCACCTTCTGAGTCTCTGCGCAACATGATGTGTGAACAG 1244
OY 721 CCCGACACATACTTCTGCTCTGCGCTCCAGGCTCATCTCTGATGATGACACCGAAC 780
DB 1245 CCCGACACATACTTCTGCTCTGCGCTCCAGGCTCATCTCTGATGATGACACCGAAC 1304
OY 781 TGCAGACATCAACGAATGTGACACAGAACCAACAGCTGCAACCTGACAGACAGTGC 840
DB 1305 TGCAGACATCAACGAATGTGACACAGAACCAACAGCTGCAACCTGACAGACAGTGC 1364
OY 841 TACAAATTACAGGGGGCTTCAATGATTCACACCCATCCGCTGTGAGAGAGCTTATCG 900
DB 1365 TACAAATTACAGGGGGCTTCAATGATTCACACCCATCCGCTGTGAGAGAGCTTATCG 1424
OY 901 AGATCAGTGAATACCGCTGATGTGTCTGCTGAGAACCCCTGGCTGACAGACAGGCC 960
DB 1425 AGATCAGTGAATACCGCTGATGTGTCTGCTGAGAACCCCTGGCTGACAGACAGGCC 1484
OY 961 TTTACATCTTGTACCGGGACATGAGCTGTGTCAGAGAGCTCCGTTCCGCTGACATC 1020
DB 1485 TTTACATCTTGTACCGGGACATGAGCTGTGTCAGAGAGCTCCGTTCCGCTGACATC 1544
OY 1021 TTCCAAATGCAAGCCAGCAGACCCGCTACCCCTGAGGCTTATTAATTTTCCAGATCAATCT 1080
DB 1545 TTCCAAATGCAAGCCAGCAGACCCGCTACCCCTGAGGCTTATTAATTTTCCAGATCAATCT 1604
OY 1081 GGGATATGAGGACAGAGATTTTACATGCGGCAAAACGGGCCCCATCAGTGCACACCTGTGTG 1140
DB 1605 GGGATATGAGGACAGAGATTTTACATGCGGCAAAACGGGCCCCATCAGTGCACACCTGTGTG 1664
OY 1141 ATGACAGCCCCCATCAAAAGGGCCCCGGGAAAATCAGCTGAGCTTGAAATGATCACTGTTC 1200
DB 1665 ATGACAGCCCCCATCAAAAGGGCCCCGGGAAAATCAGCTGAGCTTGAAATGATCACTGTTC 1724
OY 1201 AACACTGTCATCACTTCAGAGGAGCAGTCCGCTGATCCGATACCGATATATGTCGACG 1260
DB 1725 AACACTGTCATCACTTCAGAGGAGCAGTCCGCTGATCCGATACCGATATATGTCGACG 1784
OY 1261 TACCATTTC 1269
DB 1785 TACCATTTC 1793

```

RESULT 11
 AAC86968
 ID AAC86968 standard; cDNA; 2609 BP.
 AC AAC86968;
 DT 20-APR-2001 (first entry)
 XX
 DE Nucleotide sequence of human polypeptide PRO210.
 XX
 KW Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;
 KW PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;
 KW PRO1483; PRO1272; PRO419; PRO4999; PRO7170; PRO248; PRO353; PRO1318;
 KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;
 KW PRO246; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;
 KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT CDS 450..1796
 FT sig_peptide /**tag- a
 FT 450..524
 FT /**tag- b
 PN MO200077037-A2.
 PD 21-DEC-2000.

```

Db      765  ACGATCTCCAGGCGCTTATATGCCGCTTGGATACCGATGGAAGCAACCAATGT 824
QY      301  GTGATGTGACGAGTGTGCAACAGATTCCACACAGTCCAAACCCACAGATCTGCATC 360
Db      825  GTGATGTGACGAGTGTGCAACAGATTCCACACAGTCCAAACCCACAGATCTGCATC 884
QY      361  AATCTGAAGGGGGGTACCTGCTCTGACCGAGATATGGCTTCTGGAAGGCCAG 420
Db      885  AATCTGAAGGGGGGTACCTGCTCTGACCGAGATATGGCTTCTGGAAGGCCAG 944
QY      421  TGGTTAGACATTGTAATGTGCTATGTTACTGCGACGAGCTGTGCGAATGTCT 480
Db      945  TGGTTAGACATTGTAATGTGCTATGTTACTGCGACGAGCTGTGCGAATGTCT 1004
QY      481  GGATCCATCTCTGTACATGCAACCTGTGTTTAACTCAATGAGATGGAAGTCTTGC 540
Db      1005  GGATCCATCTCTGTACATGCAACCTGTGTTTAACTCAATGAGATGGAAGTCTTGC 1064
Y      541  CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGGCGCAACCTGGCTCAACACTAC 600
Db      1065  CAAGATGTGAAGAGTGTGCGACCGAGAACCCCTGGCGCAACCTGGCTCAACACTAC 1124
QY      601  GGCCTCTTCACTGCTGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCTTCATTC 660
Db      1125  GGCCTCTCTCATCTGCGCTGTGACCGAGATATGAACTTGAAGAGATGGCTTCATTC 1184
QY      661  AGTATATGAGAGAGTGCACCTTCTGAGTCTCTGCTGCGCAACATGAGTGTGAACAG 720
Db      1185  AGTATATGAGAGAGTGCACCTTCTGAGTCTCTGCTGCGCAACATGAGTGTGAACAG 1244
QY      721  CCCGGACATACTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      1245  CCCGGACATACTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
QY      781  TGGCAAGACATCAACGAATGTGAGCAGACAGAACACACGTTGCAACCTGACAGACGTC 840
Db      1305  TGGCAAGACATCAACGAATGTGAGCAGACAGAACACACGTTGCAACCTGACAGACGTC 1364
QY      841  TACAAATTTACAAGGGGGCTTCAATGATGATGACACCCCTGCGCTGAGAGCCCTTACTG 900
Db      1365  TACAAATTTACAAGGGGGCTTCAATGATGATGACACCCCTGCGCTGAGAGCCCTTACTG 1424
QY      901  AGGATCAGTATACCGCTGTATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db      1425  AGGATCAGTATACCGCTGTATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
Y      961  TTACCATCTTGTACCGGGGACATGAGCGGTGCTGAGAGCGCTCCGTTCCCGCTGACATC 1020
Db      1485  TTACCATCTTGTACCGGGGACATGAGCGGTGCTGAGAGCGCTCCGTTCCCGCTGACATC 1544
QY      1021  TTCCAATGCAAGCCAGCAGACCGCTACCTGGGGCTTATTTACATTTTCCAGATCAATCT 1080
Db      1545  TTCCAATGCAAGCCAGCAGACCGCTACCTGGGGCTTATTTACATTTTCCAGATCAATCT 1604
QY      1081  GGGATGAGGGGAGAGAAATTTACATGCGGCAACCGGCCCAATCAGTGCACCTGGTG 1140
Db      1605  GGGATGAGGGGAGAGAAATTTACATGCGGCAACCGGCCCAATCAGTGCACCTGGTG 1664
QY      1141  ATGACACGCCCATCAAAAGGGGCCCGGGAATATCAGCTGAGCTTGAATGATCAGCTGTC 1200
Db      1665  ATGACACGCCCATCAAAAGGGGCCCGGGAATATCAGCTGAGCTTGAATGATCAGCTGTC 1724
QY      1201  AACACTGTCATCACTTCAAGAGGAGCTCGTATCCGACTGGAGTATATGTGTGCGAG 1260
Db      1725  AACACTGTCATCACTTCAAGAGGAGCTCGTATCCGACTGGAGTATATGTGTGCGAG 1784
QY      1261  TACCCATTC 1269
Db      1785  TACCCATTC 1793

```

RESULT 10

```

AAS46128
ID      AAS46128 standard; cDNA; 2609 bp.
XX
AC      AAS46128;
XX
DT      18-DEC-2001 (first entry)
XX
DE      Human DNA encoding PRO polypeptide sequence #204.
XX
KW      PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss;
KW      dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW      blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW      adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW      PCR primer.
XX
OS      Homo sapiens.
XX
PN      WO200168848-A2.
XX
PD      20-SEP-2001.
XX
PE      28-FEB-2001; 2001WO-US06520.
XX
PR      01-MAR-2000; 2000WO-US05601.
PR      02-MAR-2000; 2000WO-US05841.
PR      03-MAR-2000; 2000US-187202P.
PR      06-MAR-2000; 2000US-186968P.
PR      14-MAR-2000; 2000US-189320P.
PR      15-MAR-2000; 2000WO-US06884.
PR      21-MAR-2000; 2000US-190828P.
PR      21-MAR-2000; 2000US-191007P.
PR      21-MAR-2000; 2000US-191048P.
PR      21-MAR-2000; 2000US-191314P.
PR      28-MAR-2000; 2000US-192655P.
PR      29-MAR-2000; 2000US-193032P.
PR      30-MAR-2000; 2000WO-US08439.
PR      04-APR-2000; 2000US-194449P.
PR      04-APR-2000; 2000US-194647P.
PR      11-APR-2000; 2000US-195975P.
PR      11-APR-2000; 2000US-196000P.
PR      11-APR-2000; 2000US-196590P.
PR      11-APR-2000; 2000US-196820P.
PR      18-APR-2000; 2000US-198121P.
PR      18-APR-2000; 2000US-198585P.
PR      25-APR-2000; 2000US-199397P.
PR      25-APR-2000; 2000US-199550P.
PR      25-APR-2000; 2000US-199654P.
PR      03-MAY-2000; 2000US-201516P.
PR      17-MAY-2000; 2000WO-US13705.
PR      22-MAY-2000; 2000WO-US14042.
PR      30-MAY-2000; 2000WO-US14941.
PR      02-JUN-2000; 2000WO-US15264.
PR      05-JUN-2000; 2000US-209832P.
PR      28-JUL-2000; 2000WO-US20710.
PR      22-AUG-2000; 2000US-0644848.
PR      24-AUG-2000; 2000WO-US23328.
PR      08-NOV-2000; 2000WO-US30952.
PR      01-DEC-2000; 2000WO-US32678.
PR      20-DEC-2000; 2000WO-US34956.
XX
PA      (GETH ) GENENTECH INC.
XX
PI      Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI      Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX      WPI: 2001-602746/68.
XX      P-PSDB; AAU29227.
XX
PT      Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT      presence of tumours, such as prostate and breast tumours, in mammals and

```

```

QY 781 TCCAGACATCAAGATGTGACACAGAACACAGTGCACCTGCAGACAGCTGC 840
Db 1237 TCCAGACATCAAGATGTGACACAGAACACAGTGCACCTGCAGACAGCTGC 1296
QY 841 TACAATTTACAAAGGGGCTTCAAAATGCATGCAGACCCCATCCGCTGTAGAGAGCTTATCTG 900
Db 1297 TACAATTTACAAAGGGGCTTCAAAATGCATGCAGACCCCATCCGCTGTAGAGAGCTTATCTG 1356
QY 901 AGGATCAGTATTAACCCCTGTATGTCTCTGTAGAAACCTGGCTGCAGAACACAGCC 960
Db 1357 AGGATCAGTATTAACCCCTGTATGTCTCTGTAGAAACCTGGCTGCAGAACACAGCC 1416
QY 961 TTTACCATCTTGTACCGGACATGAGCTGTGTACAGACGCTCCGCTCCGCTGCATC 1020
Db 1417 TTTACCATCTTGTACCGGACATGAGCTGTGTACAGACGCTCCGCTCCGCTGCATC 1476
QY 1021 TTTCAATTCAGACGACGACCCGCTTACCTGGGGCTTATACATTTTCCAGATCAATCT 1080
Db 1477 TTTCAATTCAGACGACGACCCGCTTACCTGGGGCTTATACATTTTCCAGATCAATCT 1536
QY 1081 GGGAAATGAGGAGAGATTTTACATGGGGGCAAAAGGGGCCCCATGATGAGCCACCCGCTG 1140
Db 1537 GGGAAATGAGGAGAGATTTTACATGGGGGCAAAAGGGGCCCCATGATGAGCCACCCGCTG 1596
QY 1141 ATGACAGCCGCCATCAAAAGGGGCCCCGGAATCCAGCTGAGCTTGAATGATCACTGTC 1200
Db 1597 ATGACAGCCGCCATCAAAAGGGGCCCCGGAATCCAGCTGAGCTTGAATGATCACTGTC 1656
QY 1201 AACACTGTCAATCACTTCAAGAGGACGCTCGTGTATCCAGCTGCGATATATGTGCGAG 1260
Db 1657 AACACTGTCAATCACTTCAAGAGGACGCTCGTGTATCCAGCTGCGATATATGTGCGAG 1716
QY 1261 TACCATTTC 1269
Db 1717 TACCATTTC 1725

```

RESULT 9
AAK37670
ID AAK37670 standard; cDNA; 2509 BP.

AAK37670:
11-SEP-2000 (first entry)
Human EGF-like homologue (PRO217) encoded by DNA32279 cDNA.

XX Inflammatory cell infiltration; immune response; T cell proliferation;
XX anti-inflammation; anti-autoimmune; anti-diabetic; spondyloarthritis;
XX T cell-mediated disease; spondyloarthritis; sclerosis; renal disease;
XX inflammatory myopathy; hemolytic anemia; thrombocytopenia; thyroiditis;
XX diabetes mellitus; demyelinating polyneuropathy; Guillain-Barre syndrome;
XX multiple sclerosis; polyneuropathy; hepatitis; cirrhosis; enteropathy;
XX sclerosing cholangitis; inflammatory bowel disease; Whipple's disease;
XX skin disease; dermatitis; psoriasis; asthma; allergic rhinitis; tumor;
XX food hypersensitivity; urticaria; eosinophilic pneumonia; transplant;
XX idiopathic pulmonary fibrosis; graft rejection; PRO245; human; PRO217;
XX EGF-like; ss.

OS Homo sapiens.
XX
XX
XX W09914241-A2.
XX
XX
XX 25-MAR-1999.
XX
XX 17-SEP-1998; 98MO-US19437.
XX
XX 17-SEP-1997; 97US-0059119.
XX 18-SEP-1997; 97US-0059263.
XX 28-OCT-1997; 97US-0063550.
XX 12-NOV-1997; 97US-0065186.
XX 21-NOV-1997; 97US-0066364.

```

PR 24-NOV-1997; 97US-0066770.  
PR 04-JUN-1998; 98US-0088026.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Fong S, Goddard A, Gurney AL, Tumas D, Wood WI;  
XX WPI; 1999-229499/19.  
XX  
XX  
XX Composition containing novel polypeptide PRO245, its agonist or  
XX antagonist -  
XX  
XX  
XX Example 1; Fig 5A-1-5A-2; 177pp; English.  
XX  
XX This invention describes a novel composition containing (apart from a  
XX carrier or excipient), a novel PRO245 polypeptide (I), its agonist or  
XX antagonist, or their fragments, for modulating: (i) infiltration of  
XX inflammatory cells into tissue; (ii) an immune response; or (iii) T cell  
XX proliferation. The composition increases or decreases any of the effects  
XX (i)-(iii). The products of the invention have anti-inflammatory,  
XX anti-autoimmune and anti-diabetic activity. (I), and its (ant)agonists  
XX and their fragments, are used to treat immune-related diseases,  
XX particularly T cell-mediated diseases. The diseases treated include  
XX systemic lupus erythematosus, rheumatoid arthritis, juvenile chronic  
XX arthritis, spondyloarthritis, systemic sclerosis (scleroderma),  
XX idiopathic inflammatory myopathies (dermatomyositis, polymyositis),  
XX Sjogren's syndrome, systemic vasculitis, sarcoidosis, autoimmune  
XX hemolytic anemia (immune pancytopenia, paroxysmal nocturnal  
XX hemoglobinuria), autoimmune thrombocytopenia (idiopathic thrombocytopenic  
XX purpura immune-mediated thrombocytopenia), thyroiditis (Grave's disease,  
XX Hashimoto's thyroiditis, juvenile lymphocytic thyroiditis, atrophic  
XX thyroiditis), diabetes mellitus, immune-mediated renal disease  
XX (glomerulonephritis, tubulointerstitial nephritis), multiple sclerosis,  
XX idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome, chronic  
XX inflammatory demyelinating polyneuropathy, infectious hepatitis  
XX (hepatitis A, B, C, D, E and other non-hepatotropic viruses), autoimmune  
XX chronic active hepatitis, primary biliary cirrhosis, granulomatous  
XX hepatitis, and sclerosing cholangitis, inflammatory bowel disease  
XX (ulcerative colitis; Crohn's disease), gluten-sensitive enteropathy, and  
XX Whipple's disease. Autoimmune or immune-mediated skin diseases including  
XX bullous skin diseases, erythema multiforme, contact dermatitis, psoriasis,  
XX asthma, allergic rhinitis, atopic dermatitis, food hypersensitivity,  
XX urticaria, eosinophilic pneumonia, idiopathic pulmonary fibrosis,  
XX hypersensitivity pneumonitis, and transplantation associated diseases  
XX (graft rejection, and graft-versus-host-disease). (I), its (ant)agonists  
XX or fragment can also be used as an adjuvant in treatment of tumors.  
XX Antibodies against (I) can also be used for diagnosing such diseases.  
XX This sequence encodes a human EGF-like homologue (PRO217) encoded by cDNA  
XX clone DNA32279 which is described in the invention.  
XX  
XX  
XX Sequence 2509 BP; 591 A; 695 C; 590 G; 633 T; 0 other;  
XX  
XX  
XX Query Match 99.9%; Score 1267.4; DB 20; Length 2509;  
XX Best Local Similarity 99.9%; Pred. No. 0;  
XX Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX  
XX 1 CAGTGCAGATGAGCTTGTACCTGGATGCGCAGTGCAGACAGTGTATGATGATGAA 60  
Db 525 CAGTGCAGATGAGCTTGTACCTGGATGCGCAGTGCAGACAGTGTATGATGAA 584  
QY 61 TGCCGAACCATCCCGAGGCGCTGCGGAGACAGATGATGTGTTAACCAAAATGCGGG 120  
Db 585 TGCCGAACCATCCCGAGGCGCTGCGGAGAGACAGATGATGTGTTAACCAAAATGCGGG 644  
QY 121 TATTATGATCTCCCGGACAAACCTGTGTATGAGGGCGCTACCTGGAACCCCTACTCG 180  
Db 645 TATTATGATCTCCCGGACAAACCTGTGTATGAGGGCGCTACCTGGAACCCCTACTCG 704  
QY 181 ACCCCCTACTGAGTCCCTTACCAAGACAGCTGCGCCACAGCTCTCAACTCAATATGCC 240  
Db 705 ACCCCCTACTGAGTCCCTTACCAAGACAGCTGCGCCACAGCTCTCAACTCAATATGCC 764  
QY 241 ACGATCTCCAGGCGCTTATATGCGCGCTTGTGATACAGATGATGAAAGCAACATGT 300

```



```

OY 721 CCGGACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGTGATGACACCGAAC 780
DB 981 CCGGACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGTGATGACACCGAAC 1040
OY 781 TGCCAGACATCAACGATGTGACACAGAACCAACAGCTGACACCTGACAGAGAGTGC 840
DB 1041 TGCCAGACATCAACGATGTGACACAGAACCAACAGCTGACACCTGACAGAGAGTGC 1100
OY 841 TACAAATTACAGGGGCTTCAATGCAATGCAACCCATCCGCTGTGAGAGCCTTATCTG 900
DB 1101 TACAAATTACAGGGGCTTCAATGCAATGCAACCCATCCGCTGTGAGAGCCTTATCTG 1160
OY 901 AGGATGATGATACCGCTGTATGTGCTGCTGAGAACCCGCTGAGAGAGAGAGAGCC 960
DB 1161 AGGATGATGATACCGCTGTATGTGCTGCTGAGAACCCGCTGAGAGAGAGAGAGCC 1220
OY 961 TTTACCATCTTGTACCGGACATGAGAGCTGTGTCAGAGACGCTCCGCTGACATC 1020
DB 1221 TTTACCATCTTGTACCGGACATGAGAGCTGTGTCAGAGAGCTCCGCTGACATC 1280
OY 1021 TTCCAATGCAAGCGACGACCGCTGACCGCTGATTTACATTTTCAGATCAAAATCT 1080
DB 1281 TTCCAATGCAAGCGACGACCGCTGACCGCTGATTTACATTTTCAGATCAAAATCT 1340
OY 1081 GGGAAATGAGGCGAGAGAAATTTTACATGCGGCAACGGGCGCCATCAGTGGCCAGCTG 1140
DB 1341 GGGAAATGAGGCGAGAGAAATTTTACATGCGGCAACGGGCGCCATCAGTGGCTG 1400
OY 1141 ATGACACGCGCCCATCAAAAGGCGCGGAAATTCAGCTGCACTTGGAATGATCACTGTC 1200
DB 1401 ATGACACGCGCCCATCAAAAGGCGCGGAAATTCAGCTGCACTTGGAATGATCACTGTC 1460
OY 1201 AACACTGTATCAACTTCAAGAGAGAGCTGCTGATGCGATGCGGATATATGTCGAG 1260
DB 1461 AACACTGTATCAACTTCAAGAGAGAGCTGCTGATGCGATGCGGATATATGTCGAG 1520
OY 1261 TACCCATTTC 1269
DB 1521 TACCCATTTC 1529

```

RESULT 8
AAK05359
ID AAK05359 standard; DNA; 2550 BP.
XX AAK05359;
AC 07-MAY-1999 (first entry)
XX
DE Human extracellular matrix protein (ECMP)-1 encoding DNA.
XX
KM Extracellular matrix protein; ECMP-1; diagnosis; treatment; cancer;
XX Immune disorder; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 382..1728
FT /tag= a
FT /product= "ECMP-1"
XX
XX MO9900410-A2.
XX
XX 07-JAN-1999.
XX
XX 23-JUN-1998; 98MO-US13012.
XX
XX 27-JUN-1997; 97US-0884072.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Corley NC, Guegler KJ;
PI

```

XX WI: 1999-095674/08.
XX P-PSDB; AAM94281.
XX
XX New polynucleotide encoding extracellular matrix protein, ECMP-1 -
XX useful in the diagnosis, prevention and treatment of immune
XX disorders and cancer
XX
XX Claim 3: Fig 1A-G; 79pp; English.
XX
XX This DNA encodes a human extracellular matrix protein (ECMP)-1. Host
XX cells containing a vector comprising the ECMP-1 nucleic acid are used
XX for the recombinant production of the protein. ECMP-1 and its
XX (ant)agonists, are useful in the diagnosis, prevention, and treatment
XX of cancer and immune disorders.
XX
XX Sequence 2550 BP; 623 A; 690 C; 596 G; 641 T; 0 other;
XX
XX Query Match 100.0%; Score 1269; DB 20; Length 2550;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 CAGTGCAGCAATGGCTTTGACCTGATCGGACATGACAGAGCTTATGATTTGATGAA 60
DB 457 CAGTGCAGCAATGGCTTTGACCTGATCGGACATGACAGAGCTTATGATTTGATGAA 516
OY 61 TGCCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACAAATGGCGG 120
DB 517 TGCCGAACCATCCCGAGGCTGCGGAGAGACATGATGTGTAAACAAATGGCGG 576
OY 121 TATTATGATTCGCGGCAAAACCTGTGTATCGAGGCGCTTACAGAACCTTACTG 180
DB 577 TATTATGATTCGCGGCAAAACCTGTGTATCGAGGCGCTTACAGAACCTTACTG 636
OY 181 ACCCCCTACTAGTCCGATACCAAGAGAGCTGCCCCACACTGACCTCCAAACTATCC 240
DB 637 ACCCCCTACTAGTCCGATACCAAGAGAGCTGCCCCACACTGACCTCCAAACTATCC 696
OY 241 ACGATCTCCAGGCTCTTATATGCGGCTTGGATACCAAGATGATGAAACCAACTGT 300
DB 697 ACGATCTCCAGGCTCTTATATGCGGCTTGGATACCAAGATGATGAAACCAACTGT 756
OY 301 GTGATGTGAGAGTGTGCAACAGATTCGCCAGGTGACCCGACCCAGATCTGCTATC 360
DB 757 GTGATGTGAGAGTGTGCAACAGATTCGCCAGGTGACCCGACCCAGATCTGCTATC 816
OY 361 AATACCTGAAGCGGCTTACACTGCTGACCAAGAGATTTGGCTTGTGAAGCCAG 420
DB 817 AATACCTGAAGCGGCTTACACTGCTGACCAAGAGATTTGGCTTGTGAAGCCAG 876
OY 421 TGCTTAGACATTTGATGATGATGCGTATGCTTCTGACAGAGCTGTGCGAATGTCCT 480
DB 877 TGCTTAGACATTTGATGATGATGCGTATGCTTCTGACAGAGCTGTGCGAATGTCCT 936
OY 481 GGATCCATTTCTTGTACATGCAACCCGTTTACCTCAATGAGATGGAAGTCTTATTC 540
DB 937 GGATCCATTTCTTGTACATGCAACCCGTTTACCTCAATGAGATGGAAGTCTTATTC 996
OY 541 CAAGATGTGAAGAGTGTGCCACGAGAACCCCTGCTGCAAACTCGCTCAACACTAC 600
DB 997 CAAGATGTGAAGAGTGTGCCACGAGAACCCCTGCTGCAAACTCGCTCAACACTAC 1056
OY 601 GGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAGAGATGGGCTTATTC 660
DB 1057 GGCTCTTTCATCTGCGGCTGTGACCCAGATATGAACTTGAGAGATGGGCTTATTC 1116
OY 661 AGTGAATGACAGAGTGTGCTGAGTTCCTGCCAACAATGAGTGTGTAACAG 720
DB 1117 AGTGAATGACAGAGTGTGCTGAGTTCCTGCCAACAATGAGTGTGTAACAG 1176
OY 721 CCGGACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGTGATGACACCGAAC 780
DB 1177 CCGGACATATCTTGTCTGCTGCGCCCTCCAGGCTACATCTGCTGTGATGACACCGAAC 1236

```


Db 1281 TTCACAAATGACAGCCAGCAGCCCGCTACCTGGGCGCTATTACATTTTCCAGATCAAAATCT 1340
 QY 1081 GGGATGAGGGGAGAGATTTTACATGGGCGCAAGGGCCCATAGTGGCCCGCTGGTG 1140
 Db 1341 GGGATGAGGGGAGAGATTTTACATGGGCGCAAGGGCCCATAGTGGCCCGCTGGTG 1400
 QY 1141 ATGACAGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGAGCTTGAATGATCACTGTC 1200
 Db 1401 ATGACAGCGCCCATCAAAAGGGCCCGGGAATTCAGCTGAGCTTGAATGATCACTGTC 1460
 QY 1201 AACACTGTCACTCACTTCAAGAGCAGCTCCGTATCCGACATGGCGATATATGTGGCAG 1260
 Db 1461 AACACTGTCACTCACTTCAAGAGCAGCTCCGTATCCGACATGGCGATATATGTGGCAG 1520
 QY 1261 TACCATTC 1269
 Db 1521 TACCATTC 1529

RESULT 7

AA39892
 AA39892 standard; cDNA; 2362 BP.

AA39892:

21-FEB-2000 (first entry)

cDNA sequence of the human secreted protein AK647.

AK647; aortic tissue development; smooth muscle cell modulator; SCID;
 nutritional supplement; vasculogenesis; embryonic development; infection;
 cytokine activity; cell proliferation; cell differentiation; defect; HIV;
 immune deficiency; haematopoiesis regulation; tissue regrowth; diagnose;
 wound healing; restenosis; atherosclerosis; drug screen; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 185..1532
 /*tag= a
 /product= AK647
 /note= "Secreted protein"

W09960125-A2.

25-NOV-1999.

18-MAY-1999; 99MO-US10931.

19-MAY-1998; 98US-0081002.

21-MAY-1998; 98US-0083002.

(GEMV) GENETICS INST INC.

Jacobs K, McCoy JM, Racie L, Lavallie E, Treacy M, Evans C;
 Agostino M, Lu Z, Merberg D;

MPI; 2000-053298/04.
 P-PSDB; AAY57058.

Proteins, and their encoding polynucleotides, used for treating e.g.
 smooth muscle cell growth, vasculogenesis, restenosis or
 atherosclerosis

Claim 1; Page 45-46; 49pp; English.

This is the nucleotide sequence of the human secreted protein AK647. The
 polynucleotide sequence was obtained from a human foetal kidney cDNA
 library. AK647 homologues in chicks and rodents are involved in aortic
 tissue development. The spatial and temporal distribution of AK647
 indicated that it acts as an modulator of smooth muscle cells in
 vasculogenesis during embryonic development. The primary structure of
 AK647 consists of multiple EGF domains. The AK647 protein can be used as

CC a nutritional source or supplement. The protein shows both inhibitory and
 CC inducing, cytokine, cell proliferation and cell differentiation activity.
 CC The protein may also be used in the treatment of immune deficiencies and
 CC disorders, including severe combined immunodeficiency (SCID), HIV and
 CC other viral, bacterial and fungal infections. Regulation of immune
 CC responses may also be carried out by the AK647 protein. Other uses of the
 CC protein include a role in the regulation of haematopoiesis and in the
 CC treatment of myeloid and lymphoid cell deficiencies. Uses in bone,
 CC cartilage, tendon, ligament and nerve tissue regrowth are also possible,
 CC as well as for wound healing and in the treatment of ulcers and burns.
 CC The polynucleotides and proteins can be used for preventing, treating or
 CC ameliorating smooth muscle cell growth, vasculogenesis, restenosis,
 CC atherosclerosis, blood vessel remodelling and degeneration. The proteins
 CC may also have activin/inhibin, chemotactic/chemokinetic, haemostatic and
 CC thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour
 CC invasion suppressor, and tumour inhibition activity. AK647 specific
 CC antibodies can be used for promoting smooth muscle cell growth or
 CC vasculogenesis. The proteins and polynucleotides can also be used for
 CC detection, diagnosis and drug screening.

Sequence 2362 BP; 580 A; 650 C; 548 G; 584 T; 0 other;

Query Match 100.0%; Score 1269; DB 21; Length 2362;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 1269; Conservative 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGAAATGGCTTGGACCTGGATGCGCCAGTGCAGAGAGTGTATGATATGATGA 60
 Db 261 CAGTGCAGAAATGGCTTGGACCTGGATGCGCCAGTGCAGAGAGTGTATGATATGATGA 320
 QY 61 TGGCGAACCATCCCGGAGGCGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 120
 Db 321 TGGCGAACCATCCCGGAGGCGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 380
 QY 121 TATTATGACATTCGCCGAGCAAAACCTGTATGAGAGGCGCTACTGAAACCCCTACTG 180
 Db 381 TATTATGACATTCGCCGAGCAAAACCTGTATGAGAGGCGCTACTGAAACCCCTACTG 440
 QY 181 ACCCGCTACAGAGTCCGCTACCCAGACAGCTGCCAGCAGCTCAGCTCAAACTATCC 240
 Db 441 ACCCGCTACAGAGTCCGCTACCCAGACAGCTGCCAGCAGCTCAGCTCAAACTATCC 500
 QY 241 ACGATCTCAGAGCGCTTATATGCGCGCTTGGATACCAAGATGATGAAGCAACCAATGT 300
 Db 501 ACGATCTCAGAGCGCTTATATGCGCGCTTGGATACCAAGATGATGAAGCAACCAATGT 560
 QY 301 GTGATGTGACAGAGTGTGCAACAGATTCGCCAGCAGCTGCAACCCAGATCTGCATC 360
 Db 561 GTGATGTGACAGAGTGTGCAACAGATTCGCCAGCAGCTGCAACCCAGATCTGCATC 620
 QY 361 AATACGAGAGCGGCTACACCTGCTGCAACCGAGATATGTGGCTTGTGGAAGGCCAG 420
 Db 621 AATACGAGAGCGGCTACACCTGCTGCAACCGAGATATGTGGCTTGTGGAAGGCCAG 680
 QY 421 TGGTTAGACATTTGATGATGTGGTATGCTTACTGCGACAGACCTGTGCGAATGTCCT 480
 Db 681 TGGTTAGACATTTGATGATGTGGTATGCTTACTGCGACAGACCTGTGCGAATGTCCT 740
 QY 481 GGATCCATTTCTTATCATGCAACCCGTGTTTACCCTCAATGAGATGGAAGGCTTTC 540
 Db 741 GGATCCATTTCTTATCATGCAACCCGTGTTTACCCTCAATGAGATGGAAGGCTTTC 800
 QY 541 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGCAAACTGGCTCAACACCTTAC 600
 Db 801 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGCGCAAACTGGCTCAACACCTTAC 860
 QY 601 GGCTCTTCACTTCGCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCGTTCAATTC 660
 Db 861 GGCTCTTCACTTCGCGCTGTGACCCAGATATGAACTTGAGGAAGATGGCGTTCAATTC 920
 QY 661 AGTATATGACAGAGTGCAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACGAG 720
 Db 921 AGTATATGACAGAGTGCAGCTTCTGTAGTTCCTCTGCAACATGAGTGTGAACGAG 980

Query	Match	Score	1269	DB	20	Length	2362
Db	1324	GGGATGAGGGGAGAGATTTTACATGCGGCAAGCGGCCCATTCAGTGCACCCGTG	1383				
Qy	1141	ATGACAGCCCCCATCAAAAGGCCCCGGGAAATTCAGCTGACTGTGAAATGATCACTGTC	1200				
Db	1384	ATGACACCCGCCCATCAAAAGGCCCCGGGAAATTCAGCTGAGACTTGGAAATGATCACTGTC	1443				
Qy	1201	AACACTGCATCAACTTCAGAGGCGAGCTCCGATGATCCAGCTGGGATATATGTGTGCGAG	1260				
Db	1444	AACACTGTCATCAACTTCAGAGGCGAGCTCCGATGATCCAGCTGGGATATATGTGTGCGAG	1503				
Qy	1261	TACCCATTG 1269					
Db	1504	TACCCATTG 1512					
RESULT 6							
ID	AA07567	standard; cDNA; 2362 BP.					
XX	AA07567;						
DT	21-JUN-1999	(first entry)					
XX							
DE	Homo sapiens fetal kidney clone AK647 secreted protein gene.						
XX							
KW	Secreted protein; fetal kidney; ds.						
OS	Homo sapiens.						
XX							
FT	Key	Location/Qualifiers					
FT	CDS	186..1532					
FT		/*tag= a					
FT		/product= "secreted protein"					
PN	W09900405-A1.						
XX							
PD	07-JAN-1999.						
XX							
PF	29-JUN-1998;	98WO-US13530.					
XX							
PR	30-JUN-1997;	97US-0885610.					
XX							
PA	(GEMV) GENETICS INST INC.						
PI	Agostino MJ, Evans C, Jacobs K, Lavallie ER, Mccoy JM;						
PI	Merberg D, Racie LA, Treacy M;						
XX							
XX	WPI: 1999-095671/08.						
XX	P-PSDB: AAW95709.						
PT	New polynucleotides encoding secreted human proteins - are derived						
PT	from foetal kidney or adult retina cDNA libraries, used as, e.g.						
PT	potential vaccines						
XX							
PS	Disclosure; Pages 51-52; 76pp; English.						
XX							
CC	The sequence is that encoding a secreted protein from a human fetal						
CC	kidney clone AK647. Such a sequence is predicted to have biological						
CC	activities which would make them suitable for treating, preventing or						
CC	ameliorating medical conditions in humans and animals, although no						
CC	supporting data is given. Suggested activities include nutritional						
CC	activity, cytokine and cell proliferation/differentiation activity,						
CC	immune stimulating (e.g. as vaccines) or suppressing activity,						
CC	haematopoiesis regulating activity, tissue growth activity,						
CC	activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic						
CC	and thrombolytic activity, receptor/ligand activity, anti-inflammatory						
CC	activity, cadherin/tumour invasion suppressor activity, and tumour						
CC	inhibition activity. It is also stated to be useful for gene						
CC	therapy.						
XX							
SO	Sequence 2362 BP; 580 A; 650 C; 548 G; 584 T; 0 other;						

		Best Local Similarity 100.0%: Pred. No. 0: Matches 1269: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY	1	CACTGCACGAATGGCTTTTGACCTGTGATGCCACATCGACAGACAGTGTATATGATATGAA 60
Db	261	CAATGCACGAATGGCTTTTGACCTGTGATGCCACATCGACAGACAGTGTATATGATGAA 320
QY	61	TGCGGAACCATCCCGAGGCTTGCAGAGAGACATGATGTGTTTAACCAAAATGGCGG 120
Db	321	TGCGGAACCATCCCGAGGCTTGCAGAGAGACATGATGTGTTTAACCAAAATGGCGG 380
QY	121	TATTTATGCAATCCCGAGCAAAACCTGTGATTCAGAGGCGCTTACTCGAAACCCCTACTG 180
Db	381	TATTTATGCAATCCCGAGCAAAACCTGTGATTCAGAGGCGCTTACTCGAAACCCCTACTG 440
QY	181	ACCCCTACTCAGGTCGGTACCCAGCAGTGCACCACCACTCTCAGTCCAAATCTATCC 240
Db	441	ACCCCTACTCAGGTCGGTACCCAGCAGTGCACCACCACTCTCAGTCCAAATCTATCC 500
QY	241	ACGATCTCCAGGCTCTTATATGCGGTTTGGATATCCAGATGGAGTAAAGAACCAATGT 300
Db	501	ACGATCTCCAGGCTCTTATATGCGGTTTGGATATCCAGATGGAGTAAAGAACCAATGT 560
QY	301	GTGATGTGACGAGTGTGCACAGATTTCCACACAGTGCMAACCCACCACAGATGTGATC 360
Db	561	GTGATGTGACGAGTGTGCACAGATTTCCACACAGTGCMAACCCACCACAGATGTGATC 620
QY	361	AATTAATGAAGGGGGGTACACCTGCTCTCTGACACCCAGCATATTTGGCTTCTGGAAGCCAG 420
Db	621	AATTAATGAAGGGGGGTACACCTGCTCTCTGACACCCAGCATATTTGGCTTCTGGAAGCCAG 680
QY	421	TGCTTAGACATTTGATGAATGTGCGATATGTTACTGCGACAGCTCTGTGCAATGTTCCT 480
Db	681	TGCTTAGACATTTGATGAATGTGCGATATGTTACTGCGACAGCTCTGTGCAATGTTCCT 740
QY	481	GGATCCATTTCTGTGACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCTTTC 540
Db	741	GGATCCATTTCTGTGACATGCAACCCCTGTTTACCTCAATGAGATGAGAGTCTTTC 800
QY	541	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGGCAAACTCGTGTCAACACTAC 600
Db	801	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGGCAAACTCGTGTCAACACTAC 860
QY	601	GGCTTTTCATCTGCGCTGTGACCCAGAGATATGAACTTGAGGAAGATGGGCTTCATTCG 660
Db	861	GGCTTTTCATCTGCGCTGTGACCCAGAGATATGAACTTGAGGAAGATGGGCTTCATTCG 920
QY	661	AGTATATGACGAGTGTGAGCTTCTGAGATTCCTGCGCAACATGTGATGTGAAACAG 720
Db	921	AGTATATGACGAGTGTGAGCTTCTGAGATTCCTGCGCAACATGTGATGTGAAACAG 980
QY	721	CCGGGCAATATCTTCTGCTGCTGCCCTTCAGAGGCTACATCTCTGCTGATGTGCAACCGAAGC 780
Db	981	CCGGGCAATATCTTCTGCTGCTGCCCTTCAGAGGCTACATCTCTGCTGATGTGCAACCGAAGC 1040
QY	781	TGCCAACAACATGCAAGATGTGGACACAGAAACACACAGTGCACACTGCAAGCAGCTGC 840
Db	1041	TGCCAACAACATGCAAGATGTGGACACAGAAACACACAGTGCACACTGCAAGCAGCTGC 900
QY	841	TACAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCCGCTGTGAGAGCTTATCTG 960
Db	1101	TACAATTTACAAAGGGGCTTCAAAATGCAATGCAACCCATCCGCTGTGAGAGCTTATCTG 1160
QY	901	AGATATGATATTAACCGCTATATGTGTCTGCTGTGAGAACCTTGCTGTGACAGACCAAGCC 960
Db	1161	AGATATGATATTAACCGCTATATGTGTCTGCTGTGAGAACCTTGCTGTGAGAGACCAAGCC 1220
QY	961	TTTACATCTTGTACCGGAGCATGAGAGTGTGTCAAGAGCGTCCGTTCCCGCTGTGATC 1020
Db	1221	TTTACATCTTGTACCGGAGCATGAGAGTGTGTCAAGAGCGTCCGTTCCCGCTGTGATC 1280
QY	1021	TTTCAAAATGCAAGCCAGACCCGCTTACCTGGGGCTTATTAATTTCCAGATCAAAATCT 1080

QY 1141 ATGACAGCCCCCATCAAGAGGCCCCGGGAATCCAGCTGAGCTTGGAAATGATCACTGTC 1200
DB 1421 ATGACAGCCCCCATCAAGAGGCCCCGGGAATCCAGCTGAGCTTGGAAATGATCACTGTC 1480
QY 1201 AACACTGTCAATCACTTCAGAGGAGCTCCGTGATCCAGCTGGGATATATGTGTCGAG 1260
DB 1481 AACACTGTCAATCACTTCAGAGGAGCTCCGTGATCCAGCTGGGATATATGTGTCGAG 1540
QY 1261 TACCACATTC 1269
DB 1541 TACCACATTC 1549

RESULT 5
AAZ40027
ID AAZ40027 standard; DNA: 2328 BP.
AC AAZ40027;
XX 15-FEB-2000 (first entry)
DT Full length human A55 protein coding sequence.

XX A55 protein; human; smooth muscle proliferation; tissue generation;
XX vascular smooth muscle cell; arteriosclerosis; tissue repair; myoma;
XX percutaneous transluminal coronary angioplasty; blood coagulation; PTCA;
XX actin; inhibin; chemotaxis; thrombosis; cadherin; therapy;
XX tumor metastasis inhibitor; ss.
OS Homo sapiens.
XX MO9955864-A1.
XX 04-NOV-1999.
XX 28-APR-1999; 99WO-JP02284.
XX 28-APR-1998; 98JP-0119731.
XX (ONOX) ONO PHARM CO LTD.
XX Honjo T, Tashiro K, Nakamura T;
XX MPI: 2000-038647/03.
XX P-PSDB: AAY54989.
XX Novel human polypeptides for treatment of, e.g. arteriosclerosis and
XX myoma.
XX Claim 4, 5; Page 76-80; 87p; Japanese.

XX This sequence encodes the human A55 protein of the invention. The protein
XX can be used for the treatment of diseases due to abnormal proliferation
XX of smooth muscle. The polypeptides can be used according to their inhibition
XX of the proliferation of vascular smooth muscle cells, particularly in
XX treating arteriosclerosis or re-narrowing by vascular endothelial
XX thickening after percutaneous transluminal coronary angioplasty (PTCA),
XX or myoma, haemotoplastic cell-regulatory activity, cytokine activity,
XX tissue generation/repair activity, actin/inhibin activity, taxis
XX and chemotaxis activity, blood coagulation/thrombotic activity,
XX receptor/ligand activity, cadherin/tumour metastasis inhibiting activity;
XX Note: The open reading frame of this sequence, as well as the mature
XX protein coding sequence are specifically claimed in claim 4.

XX Sequence 2328 BP; 560 A; 646 C; 541 G; 581 T; 0 other:
SQ

Query Match 100.0%; Score 1269; DB 21; Length 2328;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATGCACGAATGGCTTTGACCTGGATCGCCAGTCAGACAGTGTATTAGATATTGATGAA 60

DB 244 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAGACAGTGTATTAGATATTGATGAA 303
QY 61 TGGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTTAACCAAAATGGCGG 120
DB 304 TGGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTTTAACCAAAATGGCGG 363
QY 121 TATTATGATATCCCGGACAAACCCGTGTATGAGAGGCGCTTACTCGAACCCTACTG 180
DB 364 TATTATGATATCCCGGACAAACCCGTGTATGAGAGGCGCTTACTCGAACCCTACTG 423
QY 181 ACCCCCTACTAGGCTCCGTACCCAGACAGCTGCCGCCACCACTCTCAACTATCC 240
DB 424 ACCCCCTACTAGGCTCCGTACCCAGACAGCTGCCGCCACCACTCTCAACTATCC 483
QY 241 ACGATCTCCAGGCGCTTTATATGCGCGCTTTGGATACCAAGTGAAGCAACCAATGT 300
DB 484 ACGATCTCCAGGCGCTTTATATGCGCGCTTTGGATACCAAGTGAAGCAACCAATGT 543
QY 301 GTGGATGTGAGCAGTGTGCAACAGATTCCACACAGTGCACCCACCCAGATCTGATC 360
DB 544 GTGGATGTGAGCAGTGTGCAACAGATTCCACACAGTGCACCCACCCAGATCTGATC 603
QY 361 AATACGTAAGCGGGGTACACCTGCTCTGACCCGACGAGATATGCTTCTGGAAGCGCAG 420
DB 604 AATACGTAAGCGGGGTACACCTGCTCTGACCCGACGAGATATGCTTCTGGAAGCGCAG 663
QY 421 TGCTTAACATATGATGATGTCGTATGCTTACTGTCGACAGCTCTGTGCAATGTTCT 480
DB 664 TGCTTAACATATGATGATGTCGTATGCTTACTGTCGACAGCTCTGTGCAATGTTCT 723
QY 481 GGATCTATATCTTGTACATGCAACCCGTGTTTACCTCAATGAGAGTGAAGTCTTGC 540
DB 724 GGATCTATATCTTGTACATGCAACCCGTGTTTACCTCAATGAGAGTGAAGTCTTGC 783
QY 541 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGTCGAAACCTGCTCAACACTTAC 600
DB 784 CAAGATGTGAACGAGTGTGCCACGAGAACCCCTGTCGAAACCTGCTCAACACTTAC 843
QY 601 GGCCTTTCAATCTGCGCTGTGACCCGAGATATGAACTTGAGAGATGGCGTCACTG 660
DB 844 GGCCTTTCAATCTGCGCTGTGACCCGAGATATGAACTTGAGAGATGGCGTCACTG 903
QY 661 AGTATATGACGAGTGTGAGCTTCTGTAGTCTCTGCAACATGATGTGTGAACAG 720
DB 904 AGTATATGACGAGTGTGAGCTTCTGTAGTCTCTGCAACATGATGTGTGAACAG 963
QY 721 CCCGGCACATCTTCTGCTCTGCCCTCCAGGCTACATCTGCTGGATGACAAACGAGC 780
DB 964 CCCGGCACATCTTCTGCTCTGCCCTCCAGGCTACATCTGCTGGATGACAAACGAGC 1023
QY 781 TGGCAAGACATCAACGAATGTGAGCAGACAGAACCAAGTGAACCTTGAGAGAGCTG 840
DB 1024 TGGCAAGACATCAACGAATGTGAGCAGACAGAACCAAGTGAACCTTGAGAGAGCTG 1083
QY 841 TACAAATTACAGAGGCGCTTCAAAATGATGACACCCATCCGCTGTGAGAGAGCTTATCG 900
DB 1084 TACAAATTACAGAGGCGCTTCAAAATGATGACACCCATCCGCTGTGAGAGAGCTTATCG 1143
QY 901 AGGATCAGTATACCCGCTGTATGTGCTGCTGTGAGAACCCCTGCTGAGAACCGCC 960
DB 1144 AGGATCAGTATACCCGCTGTATGTGCTGCTGTGAGAACCCCTGCTGAGAACCGCC 1203
QY 961 TTTACATCTTGTACCGGAGATGAGCTGTGAGAGAGCTGCTGCTGAGAGCTGATC 1020
DB 1204 TTTACATCTTGTACCGGAGATGAGCTGTGAGAGAGCTGCTGCTGAGAGCTGATC 1263
QY 1021 TTCCAAATGCAAGCGACAGACCCGCTACCCCTGGGCGCTATTACATTTTCCAGATCAATCT 1080
DB 1264 TTCCAAATGCAAGCGACAGACCCGCTACCCCTGGGCGCTATTACATTTTCCAGATCAATCT 1323
QY 1081 GGGATATAGGCGACAGAAATTTTACATGCGCGCAACGGGCGCCCATAGTCCACCTGCTG 1140

```

Oy      1141  ATGACAGCCCGCATTAAGAGGGCCCCGGGAAATCAGCTGTGGAATGATCACTGTC
Db      1426  ATGACAGCGCCCGCATCAAAGGGCCCCGGGAAATCAGCTGTGGAATGATCACTGTC
Oy      1201  AACACTGCATCAACTTGCAGAGCAGCTCGTGATCCGATCGGATATATGTGTGCAG
Db      1486  AACACTGCATCAACTTGCAGAGCAGCTCGTGATCCGATCGGATATATGTGTGCAG
Oy      1261  TACCATTG 1269
Db      1546  TACCATTG 1554

RESULT 4
AAK94505
ID AAK94505 standard; cDNA; 2126 BP.
XX
AC AAK94505;
XX
OT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3356.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PE 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Makematsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
PI WPI: 2001-524255/58.
XX P-PSDB: AAM93573.
XX
DR 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8: SEQ ID NO 3356; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesized by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 2126 BP; 501 A; 608 C; 520 G; 497 T; 0 other:

Query Match          100.0%; Score 1269; DB 22; Length 2126;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCAGCAATGGCTTTGACCTGGATGCCAGTCAGAGAGAGCTTTAGATATGATGAAA 60
|||||
281 CAGTGCAGCAATGGCTTTGACCTGGATGCCAGTCAGAGAGAGTGTTTAGATATGATGAAA 340

```

QY	6	TGCGCAACCATCCCGGAGCGCTGCCAGAGACATATGTGTGTAAACCAAAATGGCGGG	120
Db	341	TGCGCAACCATCCCGGAGCGCTGCCAGAGACATATGTGTGTAAACCAAAATGGCGGG	400
QY	121	TATTTATGCAATTCGCCGAGCAAAACCTGTGTATTCGAGGGCCCTTACTCGAAACCCCTACTGG	180
Db	401	TATTTATGCAATTCGCCGAGCAAAACCTGTGTATTCGAGGGCCCTTACTCGAAACCCCTACTGG	460
QY	181	ACCCCTCTCTCAGGTCGCCGATACCAGAGATGTGCCCCACCACTCTCAGCTCCAAACTATGCC	240
Db	461	ACCCCTCTCTCAGGTCGCCGATACCAGAGATGTGCCCCACCACTCTCAGCTCCAAACTATGCC	520
QY	241	ACGATCTCCAGGCGCTTATATATGCGGCTTTGGATACCAATGATGATGAAGCAACCAATGT	300
Db	521	ACGATCTCCAGGCGCTTATATATGCGGCTTTGGATACCAATGATGATGAAGCAACCAATGT	580
QY	301	GTGGATGTGGAGCAAGTGTGGCAACAGATTCCACCAAGTGCACCCCAACCCAGATCTGCATC	360
Db	581	GTGGATGTGGAGCAAGTGTGGCAACAGATTCCACCAAGTGCACCCCAACCCAGATCTGCATC	640
QY	361	AATACTGAAGCGGGGTACACCTGCTCTCGACCGAGCGGATATGTGGCTTCTGGAAAGCGCG	420
Db	641	AATACTGAAGCGGGGTACACCTGCTCTCGACCGAGCGGATATGTGGCTTCTGGAAAGCGCG	700
QY	421	TGCTTAGACATTTGATGAATGTGCGTATGTGTTACTGCCACAGCTCTGTGCGAATGTTCCT	480
Db	701	TGCTTAGACATTTGATGAATGTGCGTATGTGTTACTGCCACAGCTCTGTGCGAATGTTCCT	760
QY	481	GGATTCCTATTTCTTGACATATGCACACCCCTGGTTTTACCCTCAATGAGAGATGGAAAGCTTGC	540
Db	761	GGATTCCTATTTCTTGACATATGCACACCCCTGGTTTTACCCTCAATGAGAGATGGAAAGCTTGC	820
QY	541	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGAACAACTGGCTCAACACTTAC	600
Db	821	CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGTGAACAACTGGCTCAACACTTAC	880
QY	601	GGCTCTTTCATCTGCGCGCTGTGACCCAGATATGAACCTTGAGGAAGATGGGCTTATATGC	660
Db	881	GGCTCTTTCATCTGCGCGCTGTGACCCAGATATGAACCTTGAGGAAGATGGGCTTATATGC	940
QY	661	AGTATATGACGAGTGCAGGCTTCTGTGAGTTCTCTTCGACCAACATGATGTGTGAAACGAG	720
Db	941	AGTATATGACGAGTGCAGGCTTCTGTGAGTTCTCTTCGACCAACATGATGTGTGAAACGAG	1000
QY	721	CCCGGCAATCTCTGTGCTCTGCGCCACAGGCTAACAATCCGTGTGAGAGACCGAAGC	780
Db	1001	CCCGGCAATCTCTGTGCTCTGCGCCACAGGCTAACAATCCGTGTGAGAGACCGAAGC	1060
QY	781	TGCCAAGACATCAACGAATGTGACACAGAAACCAAGTCAACCTGTACAGACGTGC	840
Db	1061	TGCCAAGACATCAACGAATGTGACACAGAAACCAAGTCAACCTGTACAGACGTGC	1120
QY	841	TACAAATTACAAGGGGCGCTTCAATTCATGATGACCCCATCCGCTGTGAGAGCCTTATATCG	900
Db	1121	TACAAATTACAAGGGGCGCTTCAATTCATGATGACCCCATCCGCTGTGAGAGCCTTATATCG	1180
QY	901	AGGATCAGTATAACCCGTGTATGTGTCTGTGCTGAGAAACCTGTGCTGAGAGACAGGCC	960
Db	1181	AGGATCAGTATAACCCGTGTATGTGTCTGTGCTGAGAAACCTGTGCTGAGAGACAGGCC	1240
QY	961	TTTATACATCTTTATACGGGACATGAGACGTGGTGTCAAGAGCGCTCGTGTCCCGCTGACATC	1020
Db	1241	TTTATACATCTTTATACGGGACATGAGACGTGGTGTCAAGAGCGCTCGTGTCCCGCTGACATC	1300
QY	1021	TTTCCAAATGCAACGACGACCGGTATCCGAGGGCGCTTATACAAATTTCCAGATCAAAATCT	1080
Db	1301	TTTCCAAATGCAACGACGACCGGTATCCGAGGGCGCTTATACAAATTTTCCAGATCAAAATCT	1360
QY	1081	GGGAATGAGGGCGAGAGATTTTACATGCGGCAAAACGGGCCCATCATCACTGTGCACCCCTGTGT	1140
Db	1361	GGGAATGAGGGCGAGAGATTTTACATGCGGCAAAACGGGCCCATCATCACTGTGCACCCCTGTGT	1420

XX	Homo sapiens.
OS	
XX	
XX	Key
FT	Location/Qualifiers
FT	CDS
FT	211..1557
FT	/tag= a
FT	/product= "EEGF"
FT	/note= "This CDS minus the stop codon (211-2554) is specifically claimed in claim 5"
FT	sig_peptide
FT	211..285
FT	/tag= b
FT	286..1554
FT	/tag= c
FT	/label= Mature_EEGF
FT	/note= "this region is specifically claimed in claim 4"
PN	
XX	US2001051358-A1.
PD	
XX	13-DEC-2001.
PF	
XX	25-MAR-1999; 99US-0275805.
PR	
XX	11-APR-1997; 97US-0839525.
XX	10-APR-1996; 96WO-US05247.
PA	(OLSE/) OLSEN H S.
PA	(LITH/) LI H.
XX	
XX	Olsen HS, LI H;
PI	
DR	WPI; 2002-121417/16.
DR	P-PSDB; MAU75494.
XX	
PT	New nucleic acid encoding human extracellular/epidermal growth factor,
PT	useful for diagnosis and treatment of e.g. Marfan syndrome and wounds,
PT	also related polypeptides
XX	
XX	Claim 4; Fig 1; 22pp; English.
PS	
CC	The invention relates to a novel polynucleotide which is at least 95% identical with a sequence (ATCC 97285) encoding mature human extracellular protein-like/epidermal growth factor (EGF)-like protein, EBBF. Also included are the EGGF EGF domains, a vector containing the polynucleotide, a host cell containing the vector, anti-EGGF antibodies and antagonists of EGGF. The polynucleotide is used for recombinant production of EGGF, in gene therapy, as hybridisation probes, as antisense antagonists and for chromosome identification. The protein is used to treat patients who require EGGF, to identify specific antagonists, used to treat conditions that require inhibition of EGGF (e.g. vascular smooth muscle cell proliferation, Marfan syndrome, wound healing, neurological trauma, acquired immunodeficiency syndrome (AIDS)-related dementia, ocular disorders, kidney disorders, liver disorders), hair follicle growth promotion, burns, ulcers, corneal incisions, corneal inflammation, neoplasms and psoriasis), to raise CC specific antibodies and to characterise receptors. The present CC sequence encodes EBBF.
SO	Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other;
Yy	Query Match 100.0%; Score 1269; DB 24; Length 1720; Best Local Similarity 100.0%; Pred. No. 0; Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0
1	CAGTGCACGAATGCGCTTGACCTGGATGCCAGTCAGCAGTGTGATTGATGAA 60

Db	286	AAGTGCAGCATGGCGTTTGACCTGGATGCCAGTCCAGTCAAGTGTATTGATTTGATGA	345
QY	61	TGCCGAACCATCCCGAGGCGCTGCCGAGAGACATGATGTGTATTACCAAAATGGCGGG	120
Db	346	TGCCGAACCAATCCCGAGGCGCTGCCGAGAGACATGATGTGTATTACCAAAATGGCGGG	405
QY	121	TATTATGCAATCCCGGACCAACCCGTGTATGAGGGGCCCTACTGCACCCCTACTGC	180
Db	406	TATTATGCAATCCCGGACCAACCCGTGTATGAGGGGCCCTACTGCACCCCTACTGC	465
QY	181	ACCCCTACTAGGTCCGTACCCAGACAGCTCCCCACACTCTCAGCTCCAAACTATCC	240
Db	466	ACCCCTACTAGGTCCGTACCCAGACAGCTCCCCACACTCTCAGCTCCAAACTATCC	525
QY	241	ACGATCTCCAGGCGCTTTATATATGCCCGCTTTGGATACCAATGATGAAAGCAACCAATGT	300
Db	526	ACGATCTCCAGGCGCTTTATATATGCCCGCTTTGGATACCAATGATGAAAGCAACCAATGT	585
QY	301	GTGATGTGGACAGAGTGTGCAACAGATTTCCACCAAGTGCACCCCAACCAATGTGCATC	360
Db	586	GTGATGTGGACAGAGTGTGCAACAGATTTCCACCAAGTGCACCCCAACCAATGTGCATC	645
QY	361	AATCTGAAGCGGGTATACCTGCTCCGACCGAGATTTGGCTCTGGAAGCGCAG	420
Db	646	AATCTGAAGCGGGTATACCTGCTCCGACCGAGATTTGGCTCTGGAAGCGCAG	705
QY	421	TGCTTAGACATTTGATGATGTGCGTATGGTATCTGCGACAGCTCTGTGCAATGTCTCT	480
Db	706	TGCTTAGACATTTGATGATGTGCGTATGGTATCTGCGACAGCTCTGTGCGAATGTCTCT	765
QY	481	GGATCTTATCTTGTACATGCAACCTGTGTTTACCTCAATGAGGATGAAAGTCTTGC	540
Db	766	GGATCTTATCTTGTACATGCAACCTGTGTTTACCTCAATGAGGATGAAAGTCTTGC	825
QY	541	CAAGTGTGAACGAGTGTGCCACGAGAACCCGTGCGCAACCTGTGCTAACACCTAC	600
Db	826	CAAGTGTGAACGAGTGTGCCACGAGAACCCGTGCGCAACCTGTGCTAACACCTAC	885
QY	601	GGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAAGGAAGTGGCGTTTCATGC	660
Db	886	GGCTCTTTCATCTGCCGCTGTGACCCAGATATGAACCTTGAAGGAAGTGGCGTTTCATGC	945
QY	661	AGTATATGACAGAGTGCAGTTCTCTGATGTTCTCTGCGCAACATGATGTGTGAACAG	720
Db	946	AGTATATGACAGAGTGCAGTTCTCTGATGTTCTCTGCGCAACATGATGTGTGAACAG	1005
QY	721	CCCGGACATCTTCTGCTCTGCGCCTCAGAGCTACATCGTGTGATGAAACCAAGCAGC	780
Db	1006	CCCGGACATCTTCTGCTCTGCGCCTCAGAGCTACATCGTGTGATGAAACCAAGCAGC	1065
QY	781	TGCCAAGACATCAACGAATGTGACACAGAAACACAGTGCATCCTGACAGACAGCTGC	840
Db	1066	TGCCAAGACATCAACGAATGTGACACAGAAACACAGTGCATCCTGACAGACAGCTGC	1125
QY	841	TACAATTTACAGGGGGCTTCAATGCAATGCACCCCATCGCTGTGAGAGCGTTATCTG	900
Db	1126	TACAATTTACAGGGGGCTTCAATGCAATGCACCCCATCGCTGTGAGAGCGTTATCTG	1185
QY	901	AGATATGATGAACCGGTATATGTATGTGTCCTGAGAACCCGTGGCTGACAGACACAGCC	960
Db	1186	AGATATGATGAACCGGTATATGTATGTGTCCTGAGAACCCGTGGCTGACAGACACAGCC	1245
QY	961	TTTACCATCTTTGATCCGGGACATGGAGTGTGTACAGAGCGCTCGTGTCCGCTGACATC	1020
Db	1246	TTTACCATCTTTGATCCGGGACATGGAGTGTGTGTACAGAGCGCTCGTGTCCGCTGACATC	1305
QY	1021	TTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCGCTTATACATTTTCCAGATCAAAATCT	1080
Db	1306	TTTCCAAATGCAAGCCACGACCCGCTACCCCTGGGGCGCTTATACATTTTCCAGATCAAAATCT	1365
QY	1081	GGGAATAGGCGACAAGAAATTTTAAATGCGGGGAAACGGGCCCATCAGTGCACACCTGGTG	1140
Db	1366	GGGAATAGGCGACAAGAAATTTTAAATGCGGGGAAACGGGCCCATCAGTGCACACCTGGTG	1425

22-OCT-1998.
 11-APR-1997; 97MO-US06020.
 11-APR-1997; 97MO-US06020.
 (HUMA-) HUMAN GENOME SCI INC.
 LI H, Olsen HS;
 MPI: 1998-568728/48.
 P-PSDB; AAM/9739.
 New isolated extracellular/epidermal growth factor - used for regulating vascular smooth muscle cell proliferation, e.g. for enhancing neurological functions or treating neoplasia and other disorders.

Claim 1a; Fig 1A-D; 62bp; English.

This sequence encodes a novel human extracellular/epidermal growth factor-like protein, EGF. This protein can be used to regulate vascular smooth muscle cell proliferation and for restoration or enhancement of neurological functions diminished as a result of other damaging pathologies such as AIDS dementia. The protein can also be used to treat senile dementia, ocular disorders such as corneal inflammation, for targeting tumor cells, for treating kidney disorders, for liver regeneration or treating liver dysfunction, for treating wounds including all cutaneous wounds, corneal wounds, and injuries to the epithelial-lined hollow organs of the body or resulting from trauma such as burns, abrasions and cuts as well as from surgical procedures such as surgical incisions and skin grafting. The polypeptides can also be used for treating chronic conditions, such as chronic ulcers, diabetic ulcers, other non-healing (trophic) conditions, to treat Marfan syndrome, to promote hair follicular development, to stimulate growth and differentiation of various epidermal and epithelial cells in vivo and in vitro and to stimulate embryogenesis. Antagonists to EGF can be used to treat neoplasia such as cancers or tumors, skin disorders such as psoriasis or corneal inflammation. The products can also be used for identifying EGF receptors, detection, diagnosis and drug screening.

Sequence 1720 BP; 396 A; 515 C; 434 G; 372 T; 3 other;

Query Match 100.0%; Score 1269; DB 19; Length 1720;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGTGCAGCAATGGCTTTGACCTGGATGCCAGTCAGAGCAGTGTATGATATGATGA 60
 286 CAGTGCAGCAATGGCTTTGACCTGGATGCCAGTCAGAGCAGTGTATGATATGATGA 345
 61 TCCCGAAGCATCCCGAGGCTGCCGAGACATGATGTTTAACCAAAATGGCGG 120
 346 TCCCGAAGCATCCCGAGGCTGCCGAGACATGATGTTTAACCAAAATGGCGG 405
 121 TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTCTACTGAAACCTCTAC 180
 406 TATTTATGATTCCTCCGAGCAAAACCTGTATGAGAGGCTCTACTGAAACCTCTAC 465
 181 ACCCCTACTCAGGTCCGTACCAAGCAGTGCCTCCACCACTCTCAGTCAAACTATCC 240
 466 ACCCCTACTCAGGTCCGTACCAAGCAGTGCCTCCACCACTCTCAGTCAAACTATCC 525
 241 AGCATCTCAGAGCTCTTATATGCCCTTTGGATACCAAGATGATGAAGAACCAATGT 300
 536 AGCATCTCAGAGCTCTTATATGCCCTTTGGATACCAAGATGATGAAGAACCAATGT 585
 301 GTGGATGTGACAGAGTGTGAACAGATTCCACAGTGAACCCACCAAGATGTGCATC 360
 586 GTGGATGTGACAGAGTGTGAACAGATTCCACAGTGAACCCACCAAGATGTGCATC 645
 361 AATAGTGAAGCGGCTACCTGCTCTGACACGACGATATTTGGCTTCTGGAAGCGCAG 420

646 AATAGTGAAGCGGCTACCTGCTCTGACACGACGATATTTGGCTTCTGGAAGCGCAG 705
 421 TGCTTAGACATTTGATGATGATGCTGCTATGCTACTGCGACAGCTCTGCGAATGCTCT 480
 706 TGCTTAGACATTTGATGATGATGCTGCTATGCTACTGCGACAGCTCTGCGAATGCTCT 765
 481 GGATCTTATCTTGTATACATGCAACCTGTGTTTACCTTCAATGAGATGGAAGTCTTGC 540
 766 GGATCTTATCTTGTATACATGCAACCTGTGTTTACCTTCAATGAGATGGAAGTCTTGC 825
 541 CAAGATGTGAACAGATGTGCCACCGAGAACCCCTGCTGTGCAAACTGCTCAACCTTAC 600
 826 CAAGATGTGAACAGATGTGCCACCGAGAACCCCTGCTGTGCAAACTGCTCAACCTTAC 885
 601 GGCTCTTCTTCTGCGGCTGTGACCCAGATATGAACTTGAAGATGGGCTTATGTC 660
 886 GGCTCTTCTTCTGCGGCTGTGACCCAGATATGAACTTGAAGATGGGCTTATGTC 945
 661 AGTATATGACAGATGTCACCTTCTGTAGTTCTCTGCTGCAACATGATGTTGTAACGAG 720
 946 AGTATATGACAGATGTCACCTTCTGTAGTTCTCTGCTGCAACATGATGTTGTAACGAG 1005
 721 CCCGCAATACTTCTCTCTCTGCTGCTGCAAGCTATCTGCTGATGACACCGAAGC 780
 1006 CCCGCAATACTTCTCTCTCTGCTGCTGCAAGCTATCTGCTGATGACACCGAAGC 1065
 781 TGGCAAGACATCAACGAATGAGACAGAGAACACACGTGCAACCTGACAGAGCTGTC 840
 1066 TGGCAAGACATCAACGAATGAGACAGAGAACACACGTGCAACCTGACAGAGCTGTC 1125
 841 TACAATTTCAAGGGGCTTCAATGATGATGACACCCATCCCTGTGAGAGCTTATGTC 900
 1126 TACAATTTCAAGGGGCTTCAATGATGATGACACCCATCCCTGTGAGAGCTTATGTC 1185
 901 AGGATAGATATTAACCCCTGTATGTCTGCTGAGAACCTGCTGACAGACACGCC 960
 1186 AGGATAGATATTAACCCCTGTATGTCTGCTGAGAACCTGCTGACAGACACGCC 1245
 961 TTTACCATCTTGTACCGGAGCATGAGAGTGTGTGAGAGCTGCTGCTGACATGTC 1020
 1246 TTTACCATCTTGTACCGGAGCATGAGAGTGTGTGAGAGCTGCTGCTGACATGTC 1305
 1021 TTCCAATGCAACGACGACCCGCTTACCTGAGGCTTATGATTTTCCAGATCAAACTCT 1080
 1306 TTCCAATGCAACGACGACCCGCTTACCTGAGGCTTATGATTTTCCAGATCAAACTCT 1365
 1081 GGGATATGAGGGGACAGAAATTTTACATGCGGCAAAAGGCGCCATCAGTGCACCTGTG 1140
 1366 GGGATATGAGGGGACAGAAATTTTACATGCGGCAAAAGGCGCCATCAGTGCACCTGTG 1425
 1141 ATGACAGGCGCCATCAAAAGGCGCCGGAATTCAGCTGGAATGATGATCACTGTC 1200
 1426 ATGACAGGCGCCATCAAAAGGCGCCGGAATTCAGCTGGAATGATGATCACTGTC 1485
 1201 AACACTGTCACTCACTTCAAGAGGAGCTGCTGTATCCGACTGCGGATATATGTGTGCGAG 1260
 1486 AACACTGTCACTCACTTCAAGAGGAGCTGCTGTATCCGACTGCGGATATATGTGTGCGAG 1545
 1261 TACCAATTC 1269
 1546 TACCAATTC 1554

RESULT 3
 ABK13627
 ID ABK13627 standard; cDNA; 1720 BP.
 XX ABK13627;
 AC
 XX 23-APR-2002 (first entry)
 DT
 XX Human cDNA encoding extracellular protein-1-like/EGF-like protein, EGF.
 DE
 XX

PT muscle cells expressing cell proliferating growth factor and cellular
PT adherence factor
XX
PS
XX Example 2; Page 97; 98pp; English.

CC The present invention describes an artificial vascular graft (I)
CC comprising a synthetic tubular element having a luminal surface coated
CC with several endothelial cells (ECs) and/or smooth muscle cells (SMCs)
CC genetically transformed to express at least one cell proliferating growth
CC factor (CPGF) and a cellular adherence factor (CAF). (I) is useful for
CC replacing or bypassing a portion of a vascular system of an individual,
CC by implanting (I) into the vascular system of the individual, to form a
CC fluid communication between the vascular system and the graft, where the
CC synthetic tubular element of the graft is of an inner cross sectional
CC area which is substantially equivalent to an inner cross sectional area
CC of a blood vessel. The present sequence represents the human up50 cDNA
CC sequence (also known as a developmental arteries and neural crest
CC epidermal growth factor (EGF)-like protein (DANCE) gene), which is used
CC in an example from the present invention.

XX Sequence 1347 BP; 329 A; 388 C; 327 G; 303 T; 0 other;

Query Match 100.0%; Score 1269; DB 24; Length 1347;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGCAATGGCTTGAACCTGATCCAGCAGAGTGTATGATATGATGA 60
DB 76 CACTGCAGCAATGGCTTGAACCTGATCCAGCAGAGTGTATGATATGATGA 135
QY 61 TGGCGAAGCAATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAATGGCGG 120
DB 136 TGGCGAAGCAATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAATGGCGG 195
QY 121 TATTTATGCAATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAATGGCGG 180
DB 196 TATTTATGCAATCCCGAGGCTGCGAGAGAGATGATGTGTAAACCAATGGCGG 255
QY 181 ACCCCCTACTCAGGTCCGTAACCCAGCAGCTGCCACACTCTCCAGTCAACATATCC 240
DB 256 ACCCCCTACTCAGGTCCGTAACCCAGCAGCTGCCACACTCTCCAGTCAACATATCC 315
QY 241 ACATATCCAGGCTCTTATATGCGCTTTGATACCAAGATGATGAAGCAACATGT 300
DB 316 ACATATCCAGGCTCTTATATGCGCTTTGATACCAAGATGATGAAGCAACATGT 375
QY 301 GTGGATGTGAGCAGGTGCAAGCAATCCACAGAGTGAACCCACAGATCTGCATC 360
DB 376 GTGGATGTGAGCAGGTGCAAGCAATCCACAGAGTGAACCCACAGATCTGCATC 435
QY 361 AATACGTAAGGCGGATACCTGCTCCGACAGCAGATATGCTTGTGAAGCGCAG 420
DB 436 AATACGTAAGGCGGATACCTGCTCCGACAGCAGATATGCTTGTGAAGCGCAG 495
QY 421 TGGTTAGACATTTGATGATGCTGCTATGTTACTGCGACAGCTCTGTGCAATGTTCT 480
DB 496 TGGTTAGACATTTGATGATGCTGCTATGTTACTGCGACAGCTCTGTGCAATGTTCT 555
QY 481 GGATCTATCTGTATACCAACCTGGTTTACCCCAATGAGATGAGAGATCTTCC 540
DB 556 GGATCTATCTGTATACCAACCTGGTTTACCCCAATGAGATGAGAGATCTTCC 615
QY 541 CAAGATGTGAAGAGTGTGACAGCAACCCCTGCGTGAACCTGCTCAACACTTC 600
DB 616 CAAGATGTGAAGAGTGTGACAGCAACCCCTGCGTGAACCTGCTCAACACTTC 675
QY 601 GGCTCTTTATCTGCGGCTGTGACCCAGAGATATGAATTTGAGAGAAAGGCTTCAATGC 660
DB 676 GGCTCTTTATCTGCGGCTGTGACCCAGAGATATGAATTTGAGAGAAAGGCTTCAATGC 735
QY 661 AGTGAATGAGCAGAGTGTGATGATGCTGCTGCAACCAATGATGTGTAACAG 720
DB 736 AGTGAATGAGCAGAGTGTGATGATGCTGCTGCAACCAATGATGTGTAACAG 795

QY 721 CCGGCAATACCTTCTGCTCCGCTGCGCAGGCTCATCTGATGATGAACCGAGC 780
DB 796 CCGGCAATACCTTCTGCTCCGCTGCGCAGGCTCATCTGATGATGAACCGAGC 855
QY 781 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGCTGCAACCTGAGCAGAGCTGC 840
DB 856 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGCTGCAACCTGAGCAGAGCTGC 915
QY 841 TACAAATTTCAAGGGGCTTCAAAATGATCCAGCCCATCCGCTGTAGAGAGCTTATCTG 900
DB 916 TACAAATTTCAAGGGGCTTCAAAATGATCCAGCCCATCCGCTGTAGAGAGCTTATCTG 975
QY 901 AGATCAATGATCAACCCGCTGATGTCTGCTGCTGAGAACCTGAGTCAAGACAGCC 960
DB 976 AGATCAATGATCAACCCGCTGATGTCTGCTGCTGAGAACCTGAGTCAAGACAGCC 1035
QY 961 TTTACCAATCTGTGACCGGACATGAGCTGTGTGAGAGAGCTCCCTCCGCTGACATC 1020
DB 1036 TTTACCAATCTGTGACCGGACATGAGCTGTGTGAGAGAGCTCCCTCCGCTGACATC 1095
QY 1021 TTCCAATGCAAGCAGCAGACCCGCTACCCCTGGGCGCTATATATTTCCAGATCAATCT 1080
DB 1096 TTCCAATGCAAGCAGCAGACCCGCTACCCCTGGGCGCTATATATTTCCAGATCAATCT 1155
QY 1081 GGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTCCACACCTGTG 1140
DB 1156 GGAATGAGGCGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTCCACACCTGTG 1215
QY 1141 ATGACAGCCCCATCAAAAGGGGCCGGAATTCAGCTGTGAAATGATGACTGTG 1200
DB 1216 ATGACAGCCCCATCAAAAGGGGCCGGAATTCAGCTGTGAAATGATGACTGTG 1275
QY 1201 AACATGCAATCACTTACAGAGAGCTCCGCTGATCCGATATATGATGCTGAG 1260
DB 1276 AACATGCAATCACTTACAGAGAGCTCCGCTGATCCGATATATGATGCTGAG 1335
QY 1261 TACCATTTC 1269
DB 1336 TACCATTTC 1344

RESULT 2
AAV62432
ID AAV62432 standard; DNA; 1720 BP.
XX
AC AAV62432;
DT 25-JAN-1999 (first entry)
DE
XX Human EGF genomic DNA.
DE
XX Extracellular/epidermal growth factor-like protein; EGF; human; liver;
KW vascular smooth muscle cell proliferation; neurology; pathology; AIDS;
KW dementia; ocular; disorder; cornea; inflammation; tumor cell; kidney;
KW wound; epithelium; surgery; skin graft; ulcer; diabetic; trophic;
KW Marfan syndrome; treatment; hair follicle; embryogenesis; neoplasia;
KW epidermal cell; cancer; psoriasis; detection; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 211..1557
FT CDS /*tag- a
FT sig_peptide 211..285
FT FT /*tag- b
FT mat_peptide 286..1557
FT FT /*tag- c
FT FT /product= "EGF"
XX /note= "Extracellular/epidermal growth factor-like"
XX
XX W09846746-A1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 08:40:41 ; Search time 282.77 Seconds
(without alignments)
10106.379 Million cell updates/sec

Title: US-09-674-379a-15
Perfect score: 1269
Sequence: 1 caggtcagcaatgcttga.....atgttcgcagctaccatc 1269

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1269	100.0	1347	24	Human UP50 cDNA se
2	1269	100.0	1720	19	Human EGF genomic
3	1269	100.0	1720	24	Human EGF genomic
4	1269	100.0	2126	22	Human full-length
5	1269	100.0	2328	21	Human full-length
6	1269	100.0	2362	20	Human full-length
7	1269	100.0	2362	21	Human full-length
8	1269	100.0	2550	20	Human full-length
9	1267.4	99.9	2509	20	Human EGF-like hom

10	1267.4	99.9	2609	22	Human DNA encoding
11	1267.4	99.9	2609	22	Nucleotide sequenc
12	1207.2	95.1	1228	18	Human extracellular
13	1056.2	83.2	1269	21	Smooth muscle prol
14	1056.2	83.2	1269	21	Smooth muscle prol
15	1056.2	83.2	1344	21	Smooth muscle prol
16	1056.2	83.2	1383	21	Smooth muscle prol
17	1056.2	83.2	2233	21	Smooth muscle prol
18	1056.2	83.2	2233	21	Smooth muscle prol
19	1056.2	83.2	2233	21	Smooth muscle prol
20	1056.2	83.2	2429	21	Smooth muscle prol
21	843.8	66.5	1254	21	Full length mouse
22	843.8	66.5	1254	21	CDNA encoding rat
23	841.8	66.3	1254	22	CDNA encoding rat
24	776.4	61.2	778	24	Rat cDNA isolated
25	721.4	56.8	853	22	Human PRO protein
26	721.4	56.8	853	22	Human cDNA 5'-end
27	721.4	56.8	853	22	Human cDNA clone r
28	576.8	45.5	680	21	Human secreted exp
29	374.4	29.3	377	20	Human EST sequence
30	371.2	29.3	1480	21	CDNA encoding a hu
31	371.2	29.3	1525	18	Human extracellular
32	371.2	29.3	1531	20	Human cDNA sequenc
33	371.2	29.3	1707	22	Human cDNA sequenc
34	371.2	29.3	1874	21	Human cDNA sequenc
35	371.2	29.3	1875	22	Human cDNA sequenc
36	371.2	29.3	1875	22	CDNA encoding huma
37	371.2	29.3	1994	21	Lung cancer associ
38	371.2	29.3	2018	21	Human cDNA encodin
39	347.2	27.4	1513	21	CDNA encoding a p5
40	337.6	26.6	1358	21	Human cDNA encodin
41	326	25.7	354	20	Human EST sequence
42	316.6	24.9	1018	21	CDNA encoding huma
43	316.6	24.9	1018	21	Skin cell cDNA, SE
44	316.6	24.9	1018	22	Murine cDNA isolat
45	315	24.8	670	22	Human cDNA 5'-end

ALIGNMENTS

RESULT 1	
ABL52528	ABL52528 standard; cDNA; 1347 BP.
ID	ABL52528;
XX	19-JUL-2002 (first entry)
AC	Human UP50 cDNA sequence SEQ ID NO:1.
XX	
DT	Human; UP50; artificial vascular graft; cellular adherence factor;
XX	cell proliferating growth factor; DANCE; epidermal growth factor; EGF;
KW	developmental arteries and neural crest EGF-like protein; gene; ss.
KW	
XX	
OS	Homo sapiens.
XX	
PN	WO200207646-A2.
XX	
PD	31-JAN-2002.
XX	
PF	20-JUL-2001; 2001WO-IL00670.
XX	
PR	20-JUL-2000; 2000US-0620227.
XX	
PA	(MGVS-) MGVS LTD.
XX	
PI	Flugelman MY, Preis M, Gluzman Z, Koren B, Weisz A, Cohen T;
XX	WPI; 2002-361632/39.
DR	Artificial vascular graft for therapy of vascular diseases, has
XX	synthetic tubular element having a surface coated with endothelial,
PT	

Sat Jul 5 15:31:34 2003

us-09-674-379a-15.rnpb

Page 16

Search completed: July 3, 2003, 17:53:39
Job time : 207.749 secs

241 ACGATCCAGGCGCTTATATGCCGCTTGGATACGATGATGAAGCAACCATGT 300
|||||
765 ACGATCCAGGCGCTTATATGCCGCTTGGATACGATGATGAAGCAACCATGT 824
|||||
301 GTGGATGTGACGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 360
|||||
825 GTGGATGTGACGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 884
|||||
361 AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 420
|||||
885 AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 944
|||||
421 TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 480
|||||
945 TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 1004
|||||
481 GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 540
|||||
1005 GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 1064
|||||
541 CAAGATGTGAAGAGTGTGACACCGAGAACCCCTGCTGCAAACTGCTGCAACCTTAC 600
|||||
1065 CAAGATGTGAAGAGTGTGACACCGAGAACCCCTGCTGCAAACTGCTGCAACCTTAC 1124
|||||
601 GGCTCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 660
|||||
1125 GGCTCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 1184
|||||
661 AATGATGTGACGAGTGTGACACCTGCTGTGACATGCTGCTGCAACATGAGTGTGAC 720
|||||
1185 AATGATGTGACGAGTGTGACACCTGCTGTGACATGCTGCTGCAACATGAGTGTGAC 1244
|||||
721 CCCGCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
1245 CCCGCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
|||||
781 TGCCAAAGACATCAACGAATGTGACACAGGAACACACGTCGACGACGACGACGTC 840
|||||
1305 TGCCAAAGACATCAACGAATGTGACACAGGAACACACGTCGACGACGACGACGTC 1364
|||||
841 TACATTTTCAAGAGGGGCTTCAAAATGCATGACACCCATCCGCTGTGAGAGGCTTATCTG 900
|||||
1365 TACATTTTCAAGAGGGGCTTCAAAATGCATGACACCCATCCGCTGTGAGAGGCTTATCTG 1424
|||||
901 AGGATCAGTATACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
|||||
1425 AGGATCAGTATACCGCTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484
|||||
961 TTTTACATCTTGTACCGGACATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
|||||
1485 TTTTACATCTTGTACCGGACATGACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1544
|||||
1021 TTTCCAAATGCAAGCAGACGCGGCTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1080
|||||
1545 TTTCCAAATGCAAGCAGACGCGGCTACCTGGGGCTTATTAATTTTCCAGATCAATCT 1604
|||||
1081 GGGATGAGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATCAGTGCACCTGCTG 1140
|||||
1605 GGGATGAGAGGCGAGAAATTTTACATGCGGCAAAAGGCGCCCATCAGTGCACCTGCTG 1664
|||||
1141 ATGACACGCGCCCATCAAAAGGCGCCCGGGAATTCACAGTGCACCTGCTGCTGCTG 1200
|||||
1665 ATGACACGCGCCCATCAAAAGGCGCCCGGGAATTCACAGTGCACCTGCTGCTGCTG 1724
|||||
1201 AATACGTAATGCAATGTAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
|||||
1725 AATACGTAATGCAATGTAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1784
|||||
1261 TACCATTC 1269
|||||
1785 TACCATTC 1793

RESULT 13
US-10-176-757-407
; Sequence 407, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-407
Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 CAGTGCAGAAATGGCTTTGACCTGGATGCGCAGTGCAGAGTGTAGATGATGA 60
DB CAGTGCAGAAATGGCTTTGACCTGGATGCGCAGTGCAGAGTGTAGATGATGA 584
61 TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATGCGGG 120
DB TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATGCGGG 644
585 TGCCGAACATCCCGAGGCGCTGCGAGAGACATGATGTGTAAACCAATGCGGG 644
121 TATTATGATTTCCCGGCAAAACCTGTGTGATGAGGCGCTTACGAAACCTTACTG 180
DB TATTATGATTTCCCGGCAAAACCTGTGTGATGAGGCGCTTACGAAACCTTACTG 645
645 TATTATGATTTCCCGGCAAAACCTGTGTGATGAGGCGCTTACGAAACCTTACTG 704
DB TATTATGATTTCCCGGCAAAACCTGTGTGATGAGGCGCTTACGAAACCTTACTG 704
181 ACCCCCTACTAGGCTCGTACCCAGACAGTGCAGCAGCTGCACTGCAAACTATCC 240
DB ACCCCCTACTAGGCTCGTACCCAGACAGTGCAGCAGCTGCACTGCAAACTATCC 705
705 ACCCCCTACTAGGCTCGTACCCAGACAGTGCAGCAGCTGCACTGCAAACTATCC 764
DB ACCCCCTACTAGGCTCGTACCCAGACAGTGCAGCAGCTGCACTGCAAACTATCC 764
241 ACGATTCGAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAGCAACCATGT 300
DB ACGATTCGAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAGCAACCATGT 765
765 ACGATTCGAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAGCAACCATGT 824
DB ACGATTCGAGGCGCTTATATGCGGCTTTGGATACAGATGATGAAGCAACCATGT 824
301 GTGATGTGAGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 360
DB GTGATGTGAGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 825
825 GTGATGTGAGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 884
DB GTGATGTGAGAGTGTGCAACAGATTTCCACAGTCAACCCACCAGATTCGCAATC 884
361 AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 420
DB AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 885
885 AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 944
DB AATACGAAAGCGGGTACACCTGCTGTGACACGAGATTTGGCTTCTGGAAGGCCAG 944
421 TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 480
DB TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 945
945 TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 1004
DB TGCTTAGACATGTAGTAATGCGGTATGATGATGCGAGCGCTGTGCGCAATGTCTCT 1004
481 GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 540
DB GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 1005
1005 GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 1064
DB GGATCTTATCTTGTACATGCAACCCGATTTTACCTCAATGAGATGGAAGTCTTGC 1064
541 CAAGATGTGAAGAGTGTGACACCGAGAACCCCTGCTGCAAACTGCTGCAACCTTAC 600
DB CAAGATGTGAAGAGTGTGACACCGAGAACCCCTGCTGCAAACTGCTGCAACCTTAC 1065
1065 CAAGATGTGAAGAGTGTGACACCGAGAACCCCTGCTGCAAACTGCTGCAACCTTAC 1124

TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-752-407

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAGTGCAGCAATGGCTTACCTGATGCGAGTCCAGACAGTGTGATATGATGAA 60
DB 525 CAGTGCAGCAATGGCTTACCTGATGCGAGTCCAGACAGTGTGATATGATGAA 584
QY 61 TCGCGAACCATCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 120
DB 585 TCGCGAACCATCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 644
QY 121 TATTATGATTCCTCCGAGAAACCTGTGTATGAGGGCCCTTACTGAAACCCCTACTG 180
DB 645 TATTATGATTCCTCCGAGAAACCTGTGTATGAGGGCCCTTACTGAAACCCCTACTG 704
QY 181 ACCCCCTACTGAGTCCGATCCAGACAGTCCGACACTCTGACATCAAACTATGCC 240
DB 705 ACCCCCTACTGAGTCCGATCCGATCCAGACAGTCCGACACTCTGACATCAAACTATGCC 764
QY 241 ACGATCTCAGAGGCTCTTATATGCGGCTTGGATACAGATGATGAAAGCAACCAATGT 300
DB 765 ACGATCTCAGAGGCTCTTATATGCGGCTTGGATACAGATGATGAAAGCAACCAATGT 824
QY 301 GTGGATGTGACAGATGTGCACAGATTTCCACAGATGCAACCCCAACCAATGTGATC 360
DB 825 GTGGATGTGACAGATGTGCACAGATTTCCACAGATGCAACCCCAACCAATGTGATC 884
QY 361 AATACTGAAGGCGGATCACCTGCTCGACAGAGATTTGGCTTCTGGAAGGCCAG 420
DB 885 AATACTGAAGGCGGATCACCTGCTCGACAGAGATTTGGCTTCTGGAAGGCCAG 944
QY 421 TCGTTAGACATTTGATGATGATGCTGATGATGATGATGATGATGATGATGATGAT 480
DB 945 TCGTTAGACATTTGATGATGATGCTGATGATGATGATGATGATGATGATGATGAT 1004
QY 481 GGAATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 1005 GGAATCTATTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1064
QY 541 CAAGATGTGAAGAGTGTGCACAGAGAACCCCTGCGGAGCAACCTGCTGAACACTATC 600
DB 1065 CAAGATGTGAAGAGTGTGCACAGAGAACCCCTGCGGAGCAACCTGCTGAACACTATC 1124
QY 601 GGCCTCTTCAATCTGCGGCTGTGACCCAGATATGAATTTGAGGAAGATGGCGTTCAATGC 660
DB 1125 GGCCTCTTCAATCTGCGGCTGTGACCCAGATATGAATTTGAGGAAGATGGCGTTCAATGC 1184
QY 661 AAGTATGTGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 1185 AAGTATGTGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1244
QY 721 CCCGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 1245 CCCGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
QY 781 TCGCAAGACATCAAGATGTGAGCAGAGAACCAACCAACCTGACAGACAGTGC 840
DB 1305 TCGCAAGACATCAAGATGTGAGCAGAGAACCAACCAACCTGACAGACAGTGC 1364
QY 841 TACAATTTACAGAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1365 TACAATTTACAGAGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
QY 901 AAGATAGTATTAACGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 1425 AAGATAGTATTAACGCTGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
QY 961 TTTACCATCTTTGACGCGGACATGAGACGTGTGTGACAGACGCTCCGTTCCGCTGACATTC 1020

```

```

DB 1485 TTTACCATCTTTGACGCGGACATGAGACGTGTGTGACAGACGCTCCGTTCCCGCTGACATTC 1544
QY 1021 TTCCAAATGCAAGCCAGACAGCCCGCTTACCTGAGGCTTATACATTTTCCAGATCAAAATCT 1080
DB 1545 TTCCAAATGCAAGCCAGACAGCCCGCTTACCTGAGGCTTATACATTTTCCAGATCAAAATCT 1604
QY 1081 GGGAAATGAGGGGAGAGATTTTACATGCGGCAAAAGGCGCCCATCAGTCCAGCCCTGTGTG 1140
DB 1605 GGGAAATGAGGGGAGAGATTTTACATGCGGCAAAAGGCGCCCATCAGTCCAGCCCTGTGTG 1664
QY 1141 ATGACAGCCCATCAAAAGGCGCCCGGAAATTCAGTGTGATGAAATGATCACTGTG 1200
DB 1665 ATGACAGCCCATCAAAAGGCGCCCGGAAATTCAGTGTGATGAAATGATCACTGTG 1724
QY 1201 AACACTGTCAATCACTTCAAGAGAGAGTCCGCTGATCCAGTCCAGGATATATGTGTGCGAG 1260
DB 1725 AACACTGTCAATCACTTCAAGAGAGAGTCCGCTGATCCAGTCCAGGATATATGTGTGCGAG 1784
QY 1261 TACCATTC 1269
DB 1785 TACCATTC 1793

```

RESULT 12
US-10-176-482-407
Sequence 407, Application US/10176482
Publication No. US20030022296A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zhen
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C70
CURRENT APPLICATION NUMBER: US/10/176,482
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 407
LENGTH: 2609
TYPE: DNA
ORGANISM: Homo Sapien
US-10-176-482-407

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 CAGTGCAGCAATGGCTTACCTGATGCGAGTCCAGACAGTGTGATATGATGAA 60
DB 525 CAGTGCAGCAATGGCTTACCTGATGCGAGTCCAGACAGTGTGATATGATGAA 584
QY 61 TCGCGAACCATCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 120
DB 585 TCGCGAACCATCCCGAGGCTGCGAGAGACATGATGTGTAAACCAAAATGGCGG 644
QY 121 TATTATGATTCCTCCGAGAAACCTGTGTATGAGGGCCCTTACTGAAACCCCTACTG 180
DB 645 TATTATGATTCCTCCGAGAAACCTGTGTATGAGGGCCCTTACTGAAACCCCTACTG 704
QY 181 ACCCCCTACTGAGTCCGATCCAGACAGTCCGACACTCTGACATCAAACTATGCC 240
DB 705 ACCCCCTACTGAGTCCGATCCAGACAGTCCGACACTCTGACATCAAACTATGCC 764

```

Db	1125	GGCTCTCTCATGTGGCGCGCTGTGATACCCAGATATGTGACTTTGAGGAAGATGGCGTTCATATGC	1184
QY	661	AGTGATATGAGCAGAGTGCGACTTCTCTGTAGTTCCTCTGCCAACAATGACTGTGTGACACAG	720
Db	1185	AGTGATATGAGCAGAGTGCGACTTCTCTGTAGTTCCTCTGCCAACAATGACTGTGTGACACAG	1244
QY	721	CCCGGACATTAATTGCTGCTCCGCGCCCTCCAGGCTCATCTCTGTGTGTGATGACACCCGAAC	780
Db	1245	CCCGGACATTAATTGCTGCTCCGCGCCCTCCAGGCTCATCTCTGTGTGTGATGACACCCGAAC	1304
QY	781	TGCCAAGACATCAACGAATGTGAGCAGACAGAAACACAGCTCAACTGTCAGCAGACAGGCG	840
Db	1305	TGCCAAGACATCAACGAATGTGAGCAGACAGAAACACAGCTCAACTGTCAGCAGACAGGCG	1364
QY	841	TACAAATTTACAAAGGGGGGCTTCCAAATGCATGCACCCCATCGCTGTGAGGAGCCTTATCTG	900
Db	1365	TACAAATTTACAAAGGGGGGCTTCCAAATGCATGCACCCCATCGCTGTGAGGAGCCTTATCTG	1424
QY	901	AGGATCACTGATTAACCCGCTGTATGTGCTCTGTGAGAACCTGTGCTGACAGACACAGCC	960
Db	1425	AGGATCACTGATTAACCCGCTGTATGTGCTCTGTGAGAACCTGTGCTGACAGACACAGCC	1484
QY	961	TTTACCAATCTTGTACCCGGGACATGAGCGTGTGTGTGAGAGCGCTCCGTTCCCGCTGACATC	1020
Db	1485	TTTACCAATCTTGTACCCGGGACATGAGCGTGTGTGTGAGAGCGCTCCGTTCCCGCTGACATC	1544
QY	1021	TTCCCAATGCAAGCGACGACCCGGCTACCCGTGGGGCCATTAACATTTTCCAGATCAAAATCT	1080
Db	1545	TTCCCAATGCAAGCGACGACCCGGCTACCCGTGGGGCCATTAACATTTTCCAGATCAAAATCT	1604
QY	1081	GGGAATGAGGGCAGAGAAATTTTACATCGCGCAAAACGGGCCCATCACTGACGACCTGTG	1140
Db	1605	GGGAATGAGGGCAGAGAAATTTTACATCGCGCAAAACGGGCCCATCACTGACGACCTGTG	1664
QY	1141	ATGACAGCGCCCATCAAAAGGGCCCCGGGAAATCATCAGCTGAGCTTGGAAATGATCACTGTC	1200
Db	1665	ATGACAGCGCCCATCAAAAGGGCCCCGGGAAATCATCAGCTGAGCTTGGAAATGATCACTGTC	1724
QY	1201	AACACTGTCAATCACTTCAAGAGGAGCTCCGTGATCCGATCGCGGATATATGTGTGGCAG	1260
Db	1725	AACACTGTCAATCACTTCAAGAGGAGCTCCGTGATCCGATCGCGGATATATGTGTGGCAG	1784
QY	1261	TACCCATTTC	1269
Db	1785	TACCCATTTC	1793
RESULT 11			
US-10-175-752-407			
: Sequence 407, Application US/10175752			
: Publication No. US20030022295A1			
: GENERAL INFORMATION:			
: APPLICANT: Baker, Kevin P.			
: APPLICANT: Chen, Jian			
: APPLICANT: Desnoyers, Luc			
: APPLICANT: Goddard, Audrey			
: APPLICANT: Godowski, Paul J.			
: APPLICANT: Gurney, Austin L.			
: APPLICANT: Pan, James			
: APPLICANT: Smith, Victoria			
: APPLICANT: Matanabe, Colin K.			
: APPLICANT: Wood, William I.			
: APPLICANT: Zhang, Zemin			
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
: FILE REFERENCE: P3430R1C60			
: CURRENT APPLICATION NUMBER: US/10/175,752			
: CURRENT FILING DATE: 2002-06-19			
: Prior Application removed - See file wrapper or Palm			
: NUMBER OF SEQ ID NOS: 612			
: SEQ ID NO 407			
: LENGTH: 2609			

Db 1485 TTACACATCTGTACCGGAGCATGAGACGTGTGTGACAGACGCTCCGCTCCGCTGACATC 1544
Qy 1021 TTCCAAATGCAAGCAGACCGGCTACCGCTGGGGCTTATATTTTCCAGATCAATCT 1080
Db 1545 TTCCAAATGCAAGCAGACCGGCTACCGCTGGGGCTTATATTTTCCAGATCAATCT 1604
Qy 1081 GGGAAATGAGGAGAGATTTTACATGCGGCAAGAGGCGCCCATCATGATCCACCCGCTG 1140
Db 1605 GGGAAATGAGGAGAGATTTTACATGCGGCAAGAGGCGCCCATCATGATCCACCCGCTG 1664
Qy 1141 ATGACAGCCCGCATCAAGAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGTC 1200
Db 1665 ATGACAGCCCGCATCAAGAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGTC 1724
Qy 1201 AACACGTGATCACTGAGAGGAGCGCTCGGATCCGATGATATATATGTTGCGAG 1260
Db 1725 AACACGTGATCACTGAGAGGAGCGCTCGGATCCGATGATATATGTTGCGAG 1784
Qy 1261 TACCCATTTC 1269
Db 1785 TACCCATTTC 1793

RESULT 9

US-10-173-706-407
; Sequence 407, Application US/10173706
; Publication No. US2003022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhen
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-407

Query Match 99.9%: Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%: Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAGTGCAGATGCTTTGACCTGGATGCGCAGACAGTGTATATATGATGTA 60
Db 525 CAGTGCAGATGCTTTGACCTGGATGCGCAGACAGTGTATATATGATGTA 584
Qy 61 TGGCGAACAATCCCGAGGCTTGGCGAGAGACATGATGTGTAACCAAAATGGCGG 120
Db 585 TGGCGAACAATCCCGAGGCTTGGCGAGAGACATGATGTGTAACCAAAATGGCGG 644
Qy 121 TATTTATGATTTCCCGAGCAAAACCTGTGTATGAGAGGCTTACTGCAACCTTACTG 180
Db 645 TATTTATGATTTCCCGAGCAAAACCTGTGTATGAGAGGCTTACTGCAACCTTACTG 704
Qy 181 ACCCGCTACGAGTGTGCTGACGAGACAGTGTGCGGACACAGTGTGCTGCAAACTTACC 240
Db 705 ACCCGCTACGAGTGTGCTGACGAGACAGTGTGCGGACACAGTGTGCTGCAAACTTACC 764
Qy 241 ACGATCTCAGGCGCTTATATGCGGCTTGTGATACAGATGATGAAGCAACCAATGT 300

Db 765 ACGATCTCAGGCGCTTATATGCGGCTTGTGATACAGATGATGAAGCAACCAATGT 824
Qy 301 GTGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGATC 360
Db 825 GTGATGTGACAGAGTGTGCAACAGATTCACACAGTGCACCCACCCAGATCTGATC 884
Qy 361 AATAGTGAAGGCGGTACACCTGCTGCTGACAGGAGATGATGCTTGTGAGGCGAG 420
Db 885 AATAGTGAAGGCGGTACACCTGCTGCTGACAGGAGATGATGCTTGTGAGGCGAG 944
Qy 421 TGTAGATCATGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 945 TGTAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
Qy 481 GGATCTATTTCTTGTACATGCAACCTGCTTTTACCTTCATGAGATGAGAGTCTTG 540
Db 1005 GGATCTATTTCTTGTACATGCAACCTGCTTTTACCTTCATGAGATGAGAGTCTTG 1064
Qy 541 CAAGATGAGAGAGTGTGCGCAGAGAACCCCTGCGCAAACTGCGTCAACACTTAC 600
Db 1065 CAAGATGAGAGAGTGTGCGCAGAGAACCCCTGCGCAAACTGCGTCAACACTTAC 1124
Qy 601 GGCTCTTTTCACTGCGGCTGTGACCCAGATATGAATGAGAGATGAGGCTTCAATG 660
Db 1125 GGCTCTTCACTGCGGCTGTGACCCAGATATGAATGAGAGATGAGGCTTCAATG 1184
Qy 661 AGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 1185 AGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1244
Qy 721 CCGGCGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1245 CCGGCGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
Qy 781 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGTGCACAGAGAGTGC 840
Db 1305 TGGCAAGACATCAAGATGTGAGCAGAGAACCAACAGTGCACAGAGAGTGC 1364
Qy 841 TACAAATTTACAGGCGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 1365 TACAAATTTACAGGCGCTTCAAAATGATGATGATGATGATGATGATGATGATGATGAT 1424
Qy 901 AGATCAGATGATTAACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1425 AGATCAGATGATTAACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
Qy 961 TTTACATCTTGTACCGGAGATGAGCGTGTGTGAGAGCGTCCGTTCCGCTGACATC 1020
Db 1485 TTTACATCTTGTACCGGAGATGAGCGTGTGTGAGAGCGTCCGTTCCGCTGACATC 1544
Qy 1021 TTCCAAATGCAAGCAGACCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1545 TTCCAAATGCAAGCAGACCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1604
Qy 1081 GGGAAATGAGGAGAGATTTTACATGCGGCAAGAGGCGCCCATCATGATCCACCCGCTG 1140
Db 1605 GGGAAATGAGGAGAGATTTTACATGCGGCAAGAGGCGCCCATCATGATCCACCCGCTG 1664
Qy 1141 ATGACAGCCCGCATCAAGAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGTC 1200
Db 1665 ATGACAGCCCGCATCAAGAGGCGCCGGAATTCAGCTGAGCTTGAATGATCACTGTC 1724
Qy 1201 AACACTGTCACTCAAGAGGAGAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1725 AACACTGTCACTCAAGAGGAGAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1784
Qy 1261 TACCCATTTC 1269
Db 1785 TACCCATTTC 1793

RESULT 10


```
Db 765 ACGATCTCCAGGCGCTTATATGCGCTTTGGATACAGATGATGAAGAACCAACTGT 824
QY 301 GTGGATGTGACGATGTGACAGATTCACACCGTAGTCAACCCACCAATTCGCATC 360
Db 825 GTGGATGTGACGATGTGACAGATTCACACCGTAGTCAACCCACCAATTCGCATC 884
QY 361 AATAGTGAAGGCGGGTACCTGCTCTGACACCGAGATATGCTTGTGAAGCCAG 420
Db 885 AATAGTGAAGGCGGGTACCTGCTCTGACACCGAGATATGCTTGTGAAGCCAG 944
QY 421 TCGTTAAGCATGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
Db 945 TCGTTAAGCATGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1004
QY 481 GGATTCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 1005 GGATTCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAACGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 600
Db 1065 CAAGATGTGAACGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 1124
QY 601 GGCTCTTCACTCTGCGCTGTGACACCGAGATATGATGATGATGATGATGATG 660
Db 1125 GGCTCTTCACTCTGCGCTGTGACACCGAGATATGATGATGATGATGATGATG 1184
QY 661 AGTATGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 1185 AGTATGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244
QY 721 CCGGCGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 1245 CCGGCGACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
QY 781 TGGCAAGCATGACAGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 840
Db 1305 TGGCAAGCATGACAGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 1364
QY 841 TACAATTTACAAGGCGGGCTTCAATGATGATGATGATGATGATGATGATGATG 900
Db 1365 TACAATTTACAAGGCGGGCTTCAATGATGATGATGATGATGATGATGATGATG 1424
QY 901 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
Db 1425 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1484
QY 961 TTTACATCTTGTACCGGAGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 1485 TTTACATCTTGTACCGGAGATGATGATGATGATGATGATGATGATGATGATG 1544
QY 1021 TTTCAATTTACAAGGCGGGCTTCAATGATGATGATGATGATGATGATGATGATG 1080
Db 1545 TTTCAATTTACAAGGCGGGCTTCAATGATGATGATGATGATGATGATGATGATG 1604
QY 1081 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAACTGCTGCAAACTGCTGCA 1140
Db 1605 GGGAAATGAGGCGAGAGATTTTACATGCGGCAAACTGCTGCAAACTGCTGCA 1664
QY 1141 ATGACAGCGGCGCATCAAAAGGCGCGGGAATTCAGTGGAGCTTGGAAATGATC 1200
Db 1665 ATGACAGCGGCGCATCAAAAGGCGCGGGAATTCAGTGGAGCTTGGAAATGATC 1724
QY 1201 AACACTGTCACTCACTCAAGAGAGCTGCTGATCCAGTGGAGATGATGATGATG 1260
Db 1725 AACACTGTCACTCACTCAAGAGAGCTGCTGATCCAGTGGAGATGATGATGATG 1784
QY 1261 TACCCATTC 1269
Db 1785 TACCCATTC 1793
```

RESULT 7
US-10-176-758-407

```
; Sequence 407, Application US/10176758
; Publication No. US2003008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-758-407
```

Query Match 99.9% Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 CAGTGCAGATGAGCTTTGACCTGATGCGCCAGTACAGAGAGTGTATGATGATGAA 60
Db 525 CAGTGCAGATGAGCTTTGACCTGATGCGCCAGTACAGAGAGTGTATGATGATGAA 584
QY 61 TGGCAAGCATGACAGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 120
Db 585 TGGCAAGCATGACAGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 644
QY 121 TATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Db 645 TATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 704
QY 181 ACCCCCTACTAGGCTGCTTACCCAGACAGCTGCCCCAGCACTGCACTGCACTGCCC 240
Db 705 ACCCCCTACTAGGCTGCTTACCCAGACAGCTGCCCCAGCACTGCACTGCACTGCCC 764
QY 241 ACGATCTCCAGGCGCTTATATGCGGCTTTGGATACAGATGATGATGATGATGATG 300
Db 765 ACGATCTCCAGGCGCTTATATGCGGCTTTGGATACAGATGATGATGATGATGATG 824
QY 301 GTGGATGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 825 GTGGATGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 884
QY 361 AATAGTGAAGGCGGGTACCTGCTCTGACACCGAGATATGCTTGTGAAGCCAG 420
Db 885 AATAGTGAAGGCGGGTACCTGCTCTGACACCGAGATATGCTTGTGAAGCCAG 944
QY 421 TCGTTAAGCATGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
Db 945 TCGTTAAGCATGTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1004
QY 481 GGATTCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Db 1005 GGATTCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
QY 541 CAAGATGTGAACGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 600
Db 1065 CAAGATGTGAACGATGTGACACCGAGAACCCCTGCTGCAAACTGCTCAACACTAC 1124
QY 601 GGCTCTTCACTCTGCGCTGTGACACCGAGATATGATGATGATGATGATGATGATG 660
```

PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAAGAGACGTGTTATGATGATGA 60
DB |||||
OY 525 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAAGAGACGTGTTATGATGATGA 584
DB |||||
OY 61 TGCCGACACATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 120
DB |||||
OY 585 TGCCGACACATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 644
DB |||||
OY 121 TATTTATGATTCCTCCGGACAAACCCCTGTATGAGAGGGCCCTACTGCAACCCCTACTGC 180
DB |||||
OY 645 TATTTATGATTCCTCCGGACAAACCCCTGTATGAGAGGGCCCTACTGCAACCCCTACTGC 704
DB |||||
OY 181 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTCAGCTCCAACTATGCC 240
DB |||||
OY 705 ACCCCCTACTGAGTCCGTACCCAGCAGCTGCCACACACTCTCAGCTCCAACTATGCC 764
DB |||||
OY 241 AGCATCTCCAGGCTCTTATATGCCGCTTGATACCATGATGAAAGCAACCAATGT 300
DB |||||
OY 765 AGCATCTCCAGGCTCTTATATGCCGCTTGATACCATGATGAAAGCAACCAATGT 824
DB |||||
OY 301 GTGATGTGAGAGATGTGCACAGATTCACACAGTGCACACCCACCCAGATCTGCATC 360
DB |||||
OY 825 GTGATGTGAGAGATGTGCACAGATTCACACAGTGCACACCCACCCAGATCTGCATC 884
DB |||||
OY 361 AATACTGAAGGGGGGTACACCTGCTCTGCACAGCAGCAGATTTGGCTTTGGAAGGCCAG 420
DB |||||
OY 885 AATACTGAAGGGGGGTACACCTGCTCTGCACAGCAGCAGATTTGGCTTTGGAAGGCCAG 944
DB |||||
OY 421 TGCCTTAGAATGTGATGATGTGCTGATGTGATCTGCCAGACCTCTGTGGCAATGTCTT 480
DB |||||
OY 945 TGCCTTAGAATGTGATGATGTGCTGATGTGATCTGCCAGACCTCTGTGGCAATGTCTT 1004
DB |||||
OY 481 GGATCCATATCTTGTACATGCAACCCGTGTTTACCTCAATGAGATGGAAGGCTTTCG 540
DB |||||
OY 1005 GGATCCATATCTTGTACATGCAACCCGTGTTTACCTCAATGAGATGGAAGGCTTTCG 1064
DB |||||
OY 541 CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAACTCCGTCACACCTTAC 600
DB |||||
OY 1065 CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCTGCAAACTCCGTCACACCTTAC 1124
DB |||||
OY 601 GGCTCTTATATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB |||||
OY 1125 GGCTCTTATATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1184
DB |||||
OY 661 AGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB |||||
OY 1185 AGTATATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1244
DB |||||
OY 721 CCCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB |||||
OY 1245 CCCGGACATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304
DB |||||
OY 781 TGCCGAACATCAAGATGTGAGCAGAGAACCAACAGTGCACCTGCAAGCAGAGCTGC 840
DB |||||
OY 1305 TGCCGAACATCAAGATGTGAGCAGAGAACCAACAGTGCACCTGCAAGCAGAGCTGC 1364
DB |||||
OY 841 TACAAATTTACAGAGGGGGCTTCAAAATGATGACCCGCTGCTGAGAGGGCCCTTATCTG 900
DB |||||
OY 1365 TACAAATTTACAGAGGGGGCTTCAAAATGATGACCCGCTGCTGAGAGGGCCCTTATCTG 1424
DB |||||
OY 901 AGGATCAATGATTAACCGCTGTATGTGTCTGCTGAGAACCCCTGCTGCAAGACCAAGCC 960
DB |||||
OY 1425 AGGATCAATGATTAACCGCTGTATGTGTCTGCTGAGAACCCCTGCTGCAAGACCAAGCC 1484
DB |||||
OY 961 TTTTACCATCTGTACCAGGACATGACAGTGTGTGTGACAGACGCTCCGTTCCCGCTGACATC 1020
DB |||||
OY 1485 TTTTACCATCTGTACCAGGACATGACAGTGTGTGTGACAGACGCTCCGTTCCCGCTGACATC 1544
DB |||||

```

```

OY 1021 TTCCAAATGCAACCCGACGACCCGCTACCTTGCGGCTTATTACATTTTCCAGATCAATCT 1080
DB |||||
OY 1545 TTCCAAATGCAACCCGACGACCCGCTACCTTGCGGCTTATTACATTTTCCAGATCAATCT 1604
DB |||||
OY 1081 GGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCTG 1140
DB |||||
OY 1605 GGGATGAGGGGAGAGATTTTACATGCGGCAAAAGGGGCCCATCAGTGCACCCCTGCTG 1664
DB |||||
OY 1141 ATGACAGCCCATCAACAAAGGGGCCCGGAAATCCAGCTGACCTTGAATGATCACTGTC 1200
DB |||||
OY 1665 ATGACAGCCCATCAACAAAGGGGCCCGGAAATCCAGCTGACCTTGAATGATCACTGTC 1724
DB |||||
OY 1201 AACCTGTATCACTTACAGAGCAGTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1260
DB |||||
OY 1725 AACCTGTATCACTTACAGAGCAGTCCGCTGATCCGACTGCGGATATATGTGTGCGAG 1784
DB |||||
OY 1261 TACCCATTC 1269
DB |||||
OY 1785 TACCCATTC 1793
DB |||||

```

RESULT 6
US-10-174-590-407

; Sequence 407, Application US/10174590
; Publication No. US20030008352A1

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 407
; LENGTH: 2609
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-174-590-407

```

Query Match 99.9%; Score 1267.4; DB 9; Length 2609;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAAGAGACGTGTTATGATGATGA 60
DB |||||
OY 525 CAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAAGAGACGTGTTATGATGATGA 584
DB |||||
OY 61 TGCCGAACATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 120
DB |||||
OY 585 TGCCGAACATCCCGAGGCGCTGCCAGAGACATGATGTGTGTTAACCAAAATGGCGGG 644
DB |||||
OY 121 TATTTATGATTCCTCCGGACAAACCCCTGTATGAGAGGGCCCTACTGCAACCCCTACTGC 180
DB |||||
OY 645 TATTTATGATTCCTCCGGACAAACCCCTGTATGAGAGGGCCCTACTGCAACCCCTACTGC 704
DB |||||
OY 181 ACCCCCTACTGAGTCCGTACCCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB |||||
OY 705 ACCCCCTACTGAGTCCGTACCCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
DB |||||
OY 241 AGCATCTCCAGGCTCTTATATGCCGCTTGATACCATGATGAAAGCAACCAATGT 300
DB |||||

```

PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066840
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/095998
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/097000
PRIOR FILING DATE: 1998-08-18
PRIOR APPLICATION NUMBER: 60/099601
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100858
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106032
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/109304
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/139695
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: 60/145070
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/149396
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 08/918874
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 08/933821
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 08/960507
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 08/114844
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: 09/136801
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136804
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/136828
PRIOR FILING DATE: 1998-08-19
PRIOR APPLICATION NUMBER: 09/158342
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 09/202088
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: 09/254311
PRIOR FILING DATE: 1999-03-03
PRIOR APPLICATION NUMBER: 09/254460
PRIOR FILING DATE: 1999-03-09
PRIOR APPLICATION NUMBER: 09/254465
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: 09/284663

PRIOR FILING DATE: 1999-04-15
PRIOR APPLICATION NUMBER: 09/332928
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/332929
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333075
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/333077
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403296
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/423844
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 09/522342
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: 09/548815
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 09/664610
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/665350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/767609
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/808689
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/866028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/870574
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/886342
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: PCT/US98/14552
PRIOR FILING DATE: 1998-07-14
PRIOR APPLICATION NUMBER: PCT/US98/18824
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: PCT/US98/19093
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: PCT/US98/19437
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: PCT/US98/24855
PRIOR FILING DATE: 1998-11-20
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: 1998-12-01
PRIOR APPLICATION NUMBER: PCT/US98/25190
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: PCT/US99/05028
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: 1999-06-02
PRIOR APPLICATION NUMBER: PCT/US99/20111
PRIOR FILING DATE: 1999-09-01
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15

Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAGGACGAGATGGCTTTGACCTGGATGCCAGTACGACAGCTGTTAGATATGTGAA 60
Db 457 CAGGACGAGATGGCTTTGACCTGGATGCCAGTACGACAGCTGTTAGATATGTGAA 516
QY 61 TGCCGACCAATCCCGGAGGCTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG 120
Db 517 TGCGAACAATCCCGGAGGCTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG 576
QY 121 TATTTATGATTCCTCCGAGCAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTG 180
Db 577 TATTTATGATTCCTCCGAGCAACCTGTGTATCGAGGGCCCTACTCGAACCCCTACTG 636
QY 181 ACCGCTACAGGTGCGGTACCCGAGAGCTCCCGACACCTCTGAGCTCCAAATATCC 240
Db 637 ACCGCTACAGGTGCGGTACCCGAGAGCTCCCGACACCTCTGAGCTCCAAATATCC 696
QY 241 ACATCTCCAGGCTCTTATATGCGGCTTGTGATACGATGATGAAGCAACCATGT 300
Db 697 ACATCTCCAGGCTCTTATATGCGGCTTGTGATACGATGATGAAGCAACCATGT 756
QY 301 GTGATGTGACAGAGTGTGCAACGATTCACAGTGTCAACCCACCAATCTGATC 360
Db 757 GTGATGTGACAGAGTGTGCAACGATTCACAGTGTCAACCCACCAATCTGATC 816
QY 361 AATAGTGAAGGCGGTACACCTGCTGCGACGAGAGATATGCGCTCTGGAAGCCAG 420
Db 817 AATAGTGAAGGCGGTACACCTGCTGCGACGAGAGATATGCGCTCTGGAAGCCAG 876
QY 421 TGTATGACATGTGATGATGCTGCTATGCTTACTGCCAGAGCTCTGCGAATGTTCT 480
Db 877 TGTATGACATGTGATGATGCTGCTATGCTTACTGCCAGAGCTCTGCGAATGTTCT 936
QY 481 GGATCTATCTGTATGATGCAACCCGCTTTACCTCAATGAGATGAGAGTCTTGC 540
Db 937 GGATCTATCTGTATGATGCAACCCGCTTTACCTCAATGAGATGAGAGTCTTGC 996
QY 541 CAAGATGTGACAGAGTGTGCGACGAGAACCCGCTGCGCAACCTGCTCAACCTAC 600
Db 997 CAAGATGTGACAGAGTGTGCGACGAGAACCCGCTGCGCAACCTGCTCAACCTAC 1056
QY 601 GGCTCTTATCTGCGGCTGTACCCAGAGATATGATGAGAGATGAGAGTCTTGC 660
Db 1057 GGCTCTTATCTGCGGCTGTACCCAGAGATATGATGAGAGATGAGAGTCTTGC 1116
QY 661 AGTATATGAGAGTGTGAGCTCTCTGAGTCTCTGCGCAACCTGATGATGATGAG 720
Db 1117 AGTATATGAGAGTGTGAGCTCTCTGAGTCTCTGCGCAACCTGATGATGATGAG 1176
QY 721 CCCGCGACATCTTCTGCTCCTGCTCCAGGCTACATCTGCTGATGACACCGAAG 780
Db 1177 CCCGCGACATCTTCTGCTCCTGCTCCAGGCTACATCTGCTGATGACACCGAAG 1236
QY 781 TGCCGACCAATCCCGGAGGCTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG 840
Db 1237 TGCCGACCAATCCCGGAGGCTGGCGAGAGACATGATGTGTAAACCAAAATGGCGG 1296
QY 841 TACAAATTTACAAGGGGCTTCAAAATGATGACCCCATCGGCTGAGAGCCCTTACTG 900
Db 1297 TACAAATTTACAAGGGGCTTCAAAATGATGACCCCATCGGCTGAGAGCCCTTACTG 1356
QY 901 AGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAGTGTGAGAGCC 960
Db 1357 AGGATCACTGATTAACCGCTGTATGTCTGCTGAGAACCCCTGAGTGTGAGAGCC 1416
QY 961 TTTACCATCTTGTACCGGAGACATGACGCTGTCTGAGAACCCCTGAGTGTGAGAGCC 1020
Db 1417 TTTACCATCTTGTACCGGAGACATGACGCTGTCTGAGAACCCCTGAGTGTGAGAGCC 1476
QY 1021 TTCCAAATGCAAGCAGCAGCGCTACCTGGGGGCTATTTACATTTTCCAGATCAATCT 1080
Db 1477 TTCCAAATGCAAGCAGCAGCGCTACCTGGGGGCTATTTACATTTTCCAGATCAATCT 1536
```

```
QY 1081 GGAATGAGGAGCAGAGATTTTACATCGGCGCAAGGCGCCCATCATGATGCCACCTGGTG 1140
Db 1537 GGAATGAGGAGCAGAGATTTTACATCGGCGCAAGGCGCCCATCATGATGCCACCTGGTG 1596
QY 1141 ATGACAGCCCATCAAGAGGCGCCGGAATTCAGCTGTGAAATATATCATCTGTC 1200
Db 1597 ATGACAGCCCATCAAGAGGCGCCGGAATTCAGCTGTGAAATATATCATCTGTC 1656
QY 1201 AACATGTATCAACTCATGAGGAGCTCCGATCCGACAGGAGATATGATGTCGAG 1260
Db 1657 AACATGTATCAACTCATGAGGAGCTCCGATCCGACAGGAGATATGATGTCGAG 1716
QY 1261 TACCATTTC 1269
Db 1717 TACCATTTC 1725

RESULT 5
US-10-066-500-14
; Sequence 14, Application US/10066500
; Patent No. US20020177165A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Bolstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Baton
; APPLICANT: Napoleone Ferrara
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3130R1C7
; CURRENT APPLICATION NUMBER: US/10/066, 500
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002, 796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
```

US-09-083-002-1

Query Match 100.0%: Score 1269; DB 10; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGAAATGGCTTTGACCTGATGCCAGTGCAGAGAGTGTATGATGATGAA 60
 DB CAGTGCAGAAATGGCTTTGACCTGATGCCAGTGCAGAGAGTGTATGATGATGAA 320
 QY 61 TGGCGAAGCAATCCCGGAGGCGCTGGCGAGAGACATGATGTGTATACCAAAATGGCGG 120
 DB TGGCGAAGCAATCCCGGAGGCGCTGGCGAGAGACATGATGTGTATACCAAAATGGCGG 380
 QY 121 TATTTATGATCTCCCGGAGCAAAACCTGTGTATCGAGGGCCGTACTCGAACCCTACTCG 180
 DB TATTTATGATCTCCCGGAGCAAAACCTGTGTATCGAGGGCCGTACTCGAACCCTACTCG 440
 QY 181 ACCCCCTACTAGGTCCGTACCCAGACAGCTGCCCAACCACTGACCTCCAACTATCC 240
 DB ACCCCCTACTAGGTCCGTACCCAGACAGCTGCCCAACCACTGACCTCCAACTATCC 500
 QY 241 ACGATCTCCAGGCGCTTATATGCGCGCTTTGATACCAATGATGATGAAACCAATGT 300
 DB ACGATCTCCAGGCGCTTATATGCGCGCTTTGATACCAATGATGATGAAACCAATGT 560
 QY 301 GTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGCACACCCCAAGATCTGATC 360
 DB GTGATGTGAGAGAGTGTGCAACAGATTCGCCACAGTGCACACCCCAAGATCTGATC 620
 QY 361 AATAGTGAAGCGGGGTACACCTGCTCTGCACACGAGAGATGATGATGATGAAAGCGAG 420
 DB AATAGTGAAGCGGGGTACACCTGCTCTGCACACGAGAGATGATGATGATGAAAGCGAG 680
 QY 421 TGCCTTAAGCATTTGATGATGTGCTATGCTATGCTGACACAGCTCTGTGCAATGTTCT 480
 DB TGCCTTAAGCATTTGATGATGTGCTATGCTATGCTGACACAGCTCTGTGCAATGTTCT 740
 QY 481 GGATCTATTTCTTGTACATGCAACCCCTGGTTTACCCCAATGAGATGGAAGTCTGTC 540
 DB GGATCTATTTCTTGTACATGCAACCCCTGGTTTACCCCAATGAGATGGAAGTCTTTC 800
 QY 541 CAAGATGTGAAGAGTGTGCAACCGAGAACCCCTGGTGCACCAACCTGCTCAACACCTAC 600
 DB CAAGATGTGAAGAGTGTGCAACCGAGAACCCCTGGTGCACCAACCTGCTCAACACCTAC 860
 QY 601 GGCCTCTTATCTGCGCGCTGTGACCCAGAGATGAACTTGAGAGAGATGGCTTCATGTC 660
 DB GGCCTCTTATCTGCGCGCTGTGACCCAGAGATGAACTTGAGAGAGATGGCTTCATGTC 920
 QY 661 AGTATATGAGAGAGTGTGACAGCTCTGTAGTTCCTGCTGCAACATGATGATGGAACCGAG 720
 DB AGTATATGAGAGAGTGTGACAGCTCTGTAGTTCCTGCTGCAACATGATGATGGAACCGAG 980
 QY 721 CCCGCAATATCTTGTCTGCTGCGCTCCAGGCTACATCTCTGTGATGACCAACCGAGC 780
 DB CCCGCAATATCTTGTCTGCTGCGCTCCAGGCTACATCTCTGTGATGACCAACCGAGC 1040
 QY 781 TGGCAAGACATCAAGAGATGTGAGACAGAGAACACACAGCTGCAAGCAGAGCTGC 840
 DB TGGCAAGACATCAAGAGATGTGAGACAGAGAACACACAGCTGCAAGCAGAGCTGC 1100
 QY 841 TACAATTTTCAAGGGGGCTTCAAAATGCAATCGACCCCATCCGCTGTAGAGAGCTTATCTG 900
 DB TACAATTTTCAAGGGGGCTTCAAAATGCAATCGACCCCATCCGCTGTAGAGAGCTTATCTG 1160
 QY 901 AGGATCAAGTAAACCGCTGTATGTCTGCTGAGAAACCTGGCTGACAGAGACAGCC 960
 DB AGGATCAAGTAAACCGCTGTATGTCTGCTGAGAAACCTGGCTGACAGAGACAGCC 1220
 QY 961 TTTTACATCTTGTACCGGAGACATGACGTGTGTAGAGACGCTCCGCTTCCCGTGCATC 1020
 DB TTTTACATCTTGTGTACCGGAGACATGACGTGTGTAGAGACGCTCCGCTTCCCGTGCATC 1280

QY 1021 TTCCAAATGCAAGCCAGCAGCCGCTACCCCTGAGGCGCTATTACATTTCCAGATCAATCT 1080
 DB TTCCAAATGCAAGCCAGCAGCCGCTACCCCTGAGGCGCTATTACATTTCCAGATCAATCT 1340
 QY 1081 GGAATGAGAGGCGAGAAATTTTACATGCGGCAAAAGCGGCCCATCATGAGTCCAGCTGTG 1140
 DB GGAATGAGAGGCGAGAAATTTTACATGCGGCAAAAGCGGCCCATCATGAGTCCAGCTGTG 1400
 QY 1141 ATGACAGCCCGCATTAAGGCGCGCGGAAATCCAGAGCTGAGCTTGAATGATCAGCTGTC 1200
 DB ATGACAGCCCGCATTAAGGCGCGCGGAAATCCAGAGCTGAGCTTGAATGATCAGCTGTC 1460
 QY 1201 AACACTGTCATCAACTTCAGAGAGCAGCTCCGTATGCGAGTGGAGATATGTGTGCGAG 1260
 DB AACACTGTCATCAACTTCAGAGAGCAGCTCCGTATGCGAGTGGAGATATGTGTGCGAG 1520
 QY 1261 TACCCATTC 1269
 DB TACCCATTC 1529

RESULT 4

US-09-836-561-2
 Sequence 2, Application US/09836561
 Patent No. US20020038006A1

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
 Corley, Neil C.

TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/836,561

FILING DATE: 16-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/212,168

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0333 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2550 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: CORNNO101

CLONE: 45517

US-09-836-561-2
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 Query Match 100.0%; Score 1269; DB 10; Length 2550;
 Best Local Similarity 100.0%; Pred. No. 0;

TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-041-016-1

Query Match 100.0%; Score 1269; DB 9; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCACGAATGCTTTGACCTGATGCGCAGTCAGACAGTGTATATATGATGA 60
 DB 261 CAGTGCACGAATGCTTTGACCTGATGCGCAGTCAGACAGTGTATATGATGA 320
 QY 61 TGCCGAACCATCCCGAGGCGCTGCGAGGAGACATGATGTGTATACCAAAATGCGGG 120
 DB 321 TGCCGAACCATCCCGAGGCGCTGCGAGGAGACATGATGTGTATACCAAAATGCGGG 380
 QY 121 TATTTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 180
 DB 381 TATTTATGCAATTCCTCCGAGCAAAACCTGTGTATCGAGGGCCCTACTGAAACCCCTACTCG 440
 QY 181 ACCCCCTACTGAGGCGCTTACCCAGACAGTGTGCGCCACCACTCTCAAGCTCAATATCCC 240
 DB 441 ACCCCCTACTGAGGCGCTTACCCAGACAGTGTGCGCCACCACTCTCAAGCTCAATATCCC 500
 QY 241 ACATCTCCAGGCGCTTATATATGCGCTTTGGATACCAAGATGATGAAGCAACCAATGT 300
 DB 501 ACATCTCCAGGCGCTTATATATGCGCTTTGGATACCAAGATGATGAAGCAACCAATGT 560
 QY 301 GTGATGTGGACGAGTGTGCAACAGATTTCCACAGTCGCAACCCACCCAGATCTGCATC 360
 DB 561 GTGATGTGGACGAGTGTGCAACAGATTTCCACAGTCGCAACCCACCCAGATCTGCATC 620
 QY 361 AATCTGAAGGCGGCTACACCTGCTCTGCAACCGAGATATGTGCTTCTGGAAGCGCAG 420
 DB 621 AATCTGAAGGCGGCTACACCTGCTCTGCAACCGAGATATGTGCTTCTGGAAGCGCAG 680
 QY 421 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTGTGCGAATGTTCTT 480
 DB 681 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTGTGCGAATGTTCTT 740
 QY 481 GGAATCCATTTCTGTACATGCAACCCGCTTTTACCCCTCAATGAGAGTGAAGGTCTTGC 540
 DB 741 GGAATCCATTTCTGTACATGCAACCCGCTTTTACCCCTCAATGAGAGTGAAGGTCTTGC 800
 QY 541 CAAGATGTGAACGAGTGTGCCACCGAGAAACCCCTGCGTGAACCTGCGTCAACCTGAC 600
 DB 801 CAAGATGTGAACGAGTGTGCCACCGAGAAACCCCTGCGTGAACCTGCGTCAACCTGAC 860
 QY 601 GGCTCTTTTCAATCTGCGCCTGTGACCCAGAGATATGAACCTTGAAGAGATGGCGTTCAATGC 660
 DB 861 GGCTCTTTTCAATCTGCGCCTGTGACCCAGAGATATGAACCTTGAAGAGATGGCGTTCAATGC 920
 QY 661 AGTATATGAGCAGAGTGTGCTTCTGAGTCTCTGCGCAACATGATGTGTGAACAG 720
 DB 921 AGTATATGAGCAGAGTGTGCTTCTGAGTCTCTGCGCAACATGATGTGTGAACAG 980
 QY 721 CCCGACACATCTTCTGCTCTGCGCTCAAGGCTACATCTCTGTGATGAGCAACCGAAGC 780
 DB 981 CCCGACACATCTTCTGCTCTGCGCTCAAGGCTACATCTCTGTGATGAGCAACCGAAGC 1040
 QY 781 TGCCAGACATCAACGAATGTGAGCAGACAGCAACCTGCAACCTGCGACAGACGTGC 840
 DB 1041 TGCCAGACATCAACGAATGTGAGCAGACAGCAACCTGCAACCTGCGACAGACGTGC 1100
 QY 841 TACAATTTACAAAGGCGCTTCAATGATGAGACCCCTGCGTGTGAGGAGCCCTATCTG 900
 DB 1101 TACAATTTACAAAGGCGCTTCAATGATGAGACCCCTGCGTGTGAGGAGCCCTATCTG 1160
 QY 901 AGGATCAGTGAATACCGCTGTATGTCTGCTGAGAACCCCTGCGTGTGAGGAGCCCTATCTG 960
 DB 1161 AGGATCAGTGAATACCGCTGTATGTCTGCTGAGAACCCCTGCGTGTGAGGAGCCCTATCTG 1220

QY 961 TTACACATCTTGTACCCGGACATGACGCTGTGTGTCAGACGCTCCGTTCCCGCTGACATC 1020
 DB 1221 TTACACATCTTGTACCCGGACATGACGCTGTGTGTCAGACGCTCCGTTCCCGCTGACATC 1280
 QY 1021 TTCCAAATGCAAGCAGACGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAAAATCT 1080
 DB 1281 TTCCAAATGCAAGCAGACGACCCGCTACCTGCGGCTTATACATTTTCCAGATCAAAATCT 1340
 QY 1081 GGGATGAGGGCAGAGATTTTACATGCGGCAACGGGCGCCATAGTGCACACCTGCTG 1140
 DB 1341 GGGATGAGGGCAGAGATTTTACATGCGGCAACGGGCGCCATAGTGCACACCTGCTG 1400
 QY 1141 ATGACACGCCCATCAAAAGGCGCCGGGAAATCCAGCTGACCTTGAATGATCACTGTG 1200
 DB 1401 ATGACACGCCCATCAAAAGGCGCCGGGAAATCCAGCTGACCTTGAATGATCACTGTG 1460
 QY 1201 AACACTGTCATCACTTCAGAGGAGCTCCCTGATCCGACCTGCGGATATATGTGTGCGAG 1260
 DB 1461 AACACTGTCATCACTTCAGAGGAGCTCCCTGATCCGACCTGCGGATATATGTGTGCGAG 1520
 QY 1261 TACCATTTC 1269
 DB 1521 TACCATTTC 1529

RESULT 3
 US-09-083-002-1
 Sequence 1, Application US/09083002
 Patent No. US20010016650A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John M.
 APPLICANT: Kacie, Lisa A.
 APPLICANT: Lavallee, Edward R.
 APPLICANT: Merberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Evans, Cheryl
 APPLICANT: Agostino, Michael
 APPLICANT: Lu, ZhiJian
 APPLICANT: Honjo, Tasuku
 APPLICANT: Tashiro, Kei
 APPLICANT: Nakamura, Tomoyuki
 TITLE OF INVENTION: SECRETED PROTEINS
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genetics Institute, Inc.
 STREET: 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: MA
 COUNTRY: U. S. A.
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,002
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sprunger, Suzanne A.
 REGISTRATION NUMBER: P-41,323
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8284
 TELEFAX: (617) 876-3851
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2362 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA

MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 211..1554
 US-09-275-805-1

Query Match 100.0%; Score 1269; DB 10; Length 1717;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGTGCAGCAATGCTTTGACCTGGATGCCAGTACAGACAGTGTATATTTGATGA 60
 DB 286 CAGTGCAGCAATGCTTTGACCTGGATGCCAGTACAGACAGTGTATATTTGATGA 345
 QY 61 TGGCGAACCAATCCCGAGAGGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 120
 DB 346 TGGCGAACCAATCCCGAGAGGCTGGCGAGAGACATGATGTGTAAACCAATGGCGGG 405
 QY 121 TATTTATGCAATCCCGAGACAAACCTGTGTATGAGGGGCCCTACTGAAACCCCTACTCG 180
 DB 406 TATTTATGCAATCCCGAGACAAACCTGTGTATGAGGGGCCCTACTGAAACCCCTACTCG 465
 QY 181 ACCCCCTACAGAGTCCGCTACCCAGACAGTCCCGACACTCTGAGCTTCAAACTATCCC 240
 DB 466 ACCCCCTACAGAGTCCGCTACCCAGACAGTCCCGACACTCTGAGCTTCAAACTATCCC 525
 QY 241 ACAGATCCAGAGCTCTTATATGCCGCTTGTGATACAGATGATGAAGAACCAATGT 300
 DB 526 ACAGATCCAGAGCTCTTATATGCCGCTTGTGATACAGATGATGAAGAACCAATGT 585
 QY 301 GTGGAGTGTGACGAGTGTGCAACAGATTCACACAGTGTGCAACCCACAGATCTGACATC 360
 DB 586 GTGGAGTGTGACGAGTGTGCAACAGATTCACACAGTGTGCAACCCACAGATCTGACATC 645
 QY 361 AATACGTAAAGGGGGGTACACTGCTCCGCAACCGAGATATGCTTGTGAAGCCAG 420
 DB 646 AATACGTAAAGGGGGGTACACTGCTCCGCAACCGAGATATGCTTGTGAAGCCAG 705
 QY 421 TGGTTAGACATTTGATGATGCTGATGATGTTACTGCCAGACAGCTGTGCGAATGTTCCT 480
 DB 706 TGGTTAGACATTTGATGATGCTGATGATGTTACTGCCAGACAGCTGTGCGAATGTTCCT 765
 QY 481 GGATCTTATTTCTGTACATGCAACCCCTGTGTTTACCTTCATGAGATGGAAGTCTTTC 540
 DB 766 GGATCTTATTTCTGTACATGCAACCCCTGTGTTTACCTTCATGAGATGGAAGTCTTTC 825
 QY 541 CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGGCAAACTGCGTCAACACCTAC 600
 DB 826 CAAGATGTGAACGAGTGTGCCACCGAGAACCCCTGCGGCAAACTGCGTCAACACCTAC 885
 QY 601 GGCCTTTTCATCTCCCGCTGTGACCCAGATATGAACTTGAGAGATGGCGTTCTATTCG 660
 DB 886 GGCCTTTTCATCTCCCGCTGTGACCCAGATATGAACTTGAGAGATGGCGTTCTATTCG 945
 QY 661 AGTGATATGACGAGTGTGAGTGTCTGAGTGTCTCTGCCAATGAGTGTGAAACCG 720
 DB 946 AGTGATATGACGAGTGTGAGTGTCTGAGTGTCTCTGCCAATGAGTGTGAAACCG 1005
 QY 721 CCGGCGACATCTCTCTCTCTGCGCTTCAGGCTACATCTCTCTGATGACACCGAAGC 780
 DB 1006 CCGGCGACATCTCTCTCTCTGCGCTTCAGGCTACATCTCTCTGATGACACCGAAGC 1065
 QY 781 TGGCAAGACATCAAGATGTGACACAGAAACCAAGTGTGCAACCTGAGAGAGAGAGTGC 840
 DB 1066 TGGCAAGACATCAAGATGTGACACAGAAACCAAGTGTGCAACCTGAGAGAGAGAGTGC 1125
 QY 841 TACAATTTACAAGGGGCTTCAAAATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCTG 900
 DB 1126 TACAATTTACAAGGGGCTTCAAAATGATGCAACCCCATCCGCTGTGAGAGAGCTTATCTG 1185
 QY 901 AGGATAGTATTAACCCCTGTATGTCTCTGTGGAACCCCTGCTGCAAGACCAAGCC 960
 DB 1186 AGGATAGTATTAACCCCTGTATGTCTCTGTGGAACCCCTGCTGCAAGACCAAGCC 1245

QY 961 TTTCACATCTGTACCGGGACATGAGCGTGTGACAGACGCTCCGCTCCGCTGACATC 1020
 DB 1246 TTTCACATCTGTACCGGGACATGAGCGTGTGACAGACGCTCCGCTCCGCTGACATC 1305
 QY 1021 TTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCTTATTTACATTTTCCAGATCAAAATCT 1080
 DB 1306 TTCCAAATGCAAGCCAGCAGCCGCTACCTGGGGCTTATTTACATTTTCCAGATCAAAATCT 1365
 QY 1081 GGGATATGAGGCAAGAAATTTTACATGTGGGCAAAAGGGGCCCATATAGTCCACCCCTGGTG 1140
 DB 1366 GGGATATGAGGCAAGAAATTTTACATGTGGGCAAAAGGGGCCCATATAGTCCACCCCTGGTG 1425
 QY 1141 ATGACAGCCCATCAAGAGGCGCCGAGAAATCCAGTGTGATGGAATGATCACTGTTC 1200
 DB 1426 ATGACAGCCCATCAAGAGGCGCCGAGAAATCCAGTGTGATGGAATGATCACTGTTC 1485
 QY 1201 AACACTGTATCAACTTCAGAGGAGCTCCGTATCCAGTCCGAGTATATGTGTGCGAG 1260
 DB 1486 AACACTGTATCAACTTCAGAGGAGCTCCGTATCCAGTCCGAGTATATGTGTGCGAG 1545
 QY 1261 TACCCATTTC 1269
 DB 1546 TACCCATTTC 1554

RESULT 2
 US-10-041-016-1
 Sequence 1, Application US/10041016
 Patent No. US20020165151A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 McCoy, John M.
 Racie, Lisa A.
 Lavalie, Edward R.
 Werberg, David
 Treacy, Maurice
 Evans, Cheryl
 Agostino, Michael
 Lu, Zhijian
 Honjo, Tasuku
 TITLE OF INVENTION: SECRETED PROTEINS
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Genetics Institute, Inc.
 STREET: 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/041,016
 FILING DATE: 07-Jan-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,002
 FILING DATE: 21-MAR-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Sprunger, Suzanne A.
 REGISTRATION NUMBER: P-41,323
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8284
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2362 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 11:41:03 ; Search time 206.749 Seconds

(without alignments)
9581.191 Million cell updates/sec

File: us-09-674-379a-15
Perfect score: 1269
Sequence: 1 cagtcgcagatgcgttga.....atgttcgacatccattc 1269

Scoring table:
Gapop 10.0, Gapext 1.0

Searched: 1085931 seqs, 780495707 residues

Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1717	US-09-275-805-1	Sequence 1, Appl
2	1269	100.0	2362	US-10-041-016-1	Sequence 1, Appl
3	1269	100.0	2362	US-09-083-002-1	Sequence 1, Appl
4	1269	100.0	2550	US-09-836-561-2	Sequence 2, Appl
5	1267.4	99.9	2609	US-10-066-500-14	Sequence 14, Appl
6	1267.4	99.9	2609	US-10-174-590-407	Sequence 407, App
7	1267.4	99.9	2609	US-10-176-758-407	Sequence 407, App
8	1267.4	99.9	2609	US-10-175-737-407	Sequence 407, App
9	1267.4	99.9	2609	US-10-173-706-407	Sequence 407, App
10	1267.4	99.9	2609	US-10-175-738-407	Sequence 407, App
11	1267.4	99.9	2609	US-10-175-752-407	Sequence 407, App
12	1267.4	99.9	2609	US-10-176-882-407	Sequence 407, App
13	1267.4	99.9	2609	US-10-176-913-407	Sequence 407, App
14	1267.4	99.9	2609	US-10-180-552-407	Sequence 407, App
15	1267.4	99.9	2609	US-10-180-557-407	Sequence 407, App
16	1267.4	99.9	2609	US-10-173-700-407	Sequence 407, App
17	1267.4	99.9	2609	US-10-174-572-407	Sequence 407, App
18	1267.4	99.9	2609	US-10-174-579-407	Sequence 407, App
19	1267.4	99.9	2609	US-10-174-579-407	Sequence 407, App

20	1267.4	99.9	2609	US-10-174-582-407	Sequence 407, App
21	1267.4	99.9	2609	US-10-174-588-407	Sequence 407, App
22	1267.4	99.9	2609	US-10-175-739-407	Sequence 407, App
23	1267.4	99.9	2609	US-10-175-740-407	Sequence 407, App
24	1267.4	99.9	2609	US-10-175-743-407	Sequence 407, App
25	1267.4	99.9	2609	US-10-176-488-407	Sequence 407, App
26	1267.4	99.9	2609	US-10-176-492-407	Sequence 407, App
27	1267.4	99.9	2609	US-10-176-747-407	Sequence 407, App
28	1267.4	99.9	2609	US-10-176-750-407	Sequence 407, App
29	1267.4	99.9	2609	US-10-176-985-407	Sequence 407, App
30	1267.4	99.9	2609	US-10-176-987-407	Sequence 407, App
31	1267.4	99.9	2609	US-10-176-991-407	Sequence 407, App
32	1267.4	99.9	2609	US-10-176-992-407	Sequence 407, App
33	1267.4	99.9	2609	US-10-176-993-407	Sequence 407, App
34	1267.4	99.9	2609	US-10-184-658-407	Sequence 407, App
35	1267.4	99.9	2609	US-10-002-796-14	Sequence 14, Appl
36	1267.4	99.9	2609	US-10-066-273-14	Sequence 14, Appl
37	1267.4	99.9	2609	US-10-066-494-14	Sequence 14, Appl
38	1267.4	99.9	2609	US-10-173-695-407	Sequence 407, App
39	1267.4	99.9	2609	US-10-173-697-407	Sequence 407, App
40	1267.4	99.9	2609	US-10-173-705-407	Sequence 407, App
41	1267.4	99.9	2609	US-10-174-576-407	Sequence 407, App
42	1267.4	99.9	2609	US-10-174-585-407	Sequence 407, App
43	1267.4	99.9	2609	US-10-174-586-407	Sequence 407, App
44	1267.4	99.9	2609	US-10-175-747-407	Sequence 407, App
45	1267.4	99.9	2609	US-10-176-481-407	Sequence 407, App

ALIGNMENTS

RESULT 1
US-09-275-805-1
Sequence 1, Application US/09275805
Patent No. US20010051358A1
GENERAL INFORMATION:
APPLICANT: OLSEN, HENRIK S.
TITLE OF INVENTION: EXTRACELLULAR EPIDERMAL GROWTH FACTOR
TITLE OF INVENTION: LIKE PROTEIN
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,525
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BROOKES, ANDERS A.
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF224
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8509
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1717 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 1..3753
PCT-US95-02251-2

Query Match 4.1%; Score 51.6; DB 5; Length 3753;
Best Local Similarity 49.1%; Pred. No. 6.8e-06;
Matches 166; Conservative 0; Mismatches 169; Indels 3; Gaps 1;

QY	466	TGTGCGAATGTTCTGATCTTATCTTGATGCAACCTGGTTTACCCCAATGAG	525
DB	2212	TGTGAGAACTTCGGGTTCTTACCGTTGCACGTGCCCCAGGGATACGAACCCGACA	2271
QY	526	GATGGAAGTCTTGCCAAAGATGTGAAGAGTGTGCCACCGAAGCCCTGC--GTGCAA	582
DB	2272	GGACGGCTCAGTTGCATAGACGTGATGACTGTGAGGCTGGAAAGTGTGCCAAGATGGC	2331
QY	583	ACCTGCGTCACACACTACGAGCTCTTCATCTGCCGCTGTGACCCAGGATATGAATGAG	642
DB	2332	ATCTGCACGACACACAGGCTCTTCCAGTGTCAAGTCCCTCCGCTATCATCTGTCA	2391
QY	643	GAAGATGCGCTTCATTCAGATGATGAGAGAGTGCAGCTTCTGTGAGTTCTCTGCCAA	702
DB	2392	AGGATCGGAGCGCTGTGAGACATGATGATGACTTCCCTCGGCTGCATCGGG	2451
QY	703	CATGAGTGTGTGAACCGCCGACATCTCTGCTCTGCCCTCCAGGCTACATCTG	762
DB	2452	GGTGAATGATCAATACCAATGTTCTCTCAGATGTCTCTGTCCTCGGCTCATCGGTTG	2511
QY	763	CTGATGACACCGAAGCTGCCAAGACATCAACGAATG	800
DB	2512	GTGGCGCGCAGAGAGTGCAGAAAGATATATGATGAGTG	2549

Search completed: July 3, 2003, 17:40:02
Job time : 59.5574 secs

QY 835 ACCTGCTACATTTACAGAGGGGCTTCAATGATGACATGCCATCCGCTGTGAGAGCCT 894
DB 3481 GAATGCAAGAACACAGAGAGGTTCTTACCAATGCCCTGTGACACAGGGGCTTCCAGGCTGTC 3540
QY 895 TATCTGAGATGATGATTAACCCGCTGTATGTCTCTGCTGAGAACCCCTGG 944
DB 3541 AATGGCACCATGTGTGAGAGCGTGAATGATGTGTGGGGAAGACCATTTG 3590

RESULT 14

PCT-US95-02251-17
Sequence 17, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentln Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 5502 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5502
PCT-US95-02251-17

Query Match 4.1%; Score 52.4; DB 5; Length 5502;
Best Local Similarity 47.4%; Pred. No. 4.8e-06;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;

QY 481 GGAATCTATTCTGTATACATGCAACCGTGTATTACCTCAATGAGAGGAGGCTTGC 540
DB 3121 GGCTCCCTTGTAGATGCTCTGTGAGCCGGGCTATGAGTACACCCAGACAAAGAGGCTGC 3180
QY 541 CAAGATGTGAACGATGCGCACGAGAGAACCCCTG---CGTCAANAACGGCTCAACACG 597
DB 3181 CGAGATGTGACGATGCGCACGAGAGGCTGTGCGCCACGCGGCTGTGCTCAACACG 3240

QY 598 TAGGCTCTTTCATCTGCGC---CTGTGACCCAGATATGAATTGAGGAAGATGGCCTT 654
DB 3241 GAGGCTCTTTCATCTGCGC---CTGTGACCCAGATATGAATTGAGGAAGATGGCCTT 3300
QY 655 CATTGAGATGATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 714
DB 3301 GCTGTGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3360
QY 715 AACCAAGCCGCGACATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
DB 3361 AATACTGAGCTCTTCTCTCTGCAAGAGCTGTGACAGAGGCTTACCGGCAACCCCTG 3420
QY 775 CGAGCTGCCAATCAACGAATGTGAGACAGACAGACAGACAGACAGACAGACAGACAG 834
DB 3421 GGCACAGATGCGAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 835 ACCTGCTACATTTACAGAGGGGCTTCAATGATGACATGCCATCCGCTGTGAGAGCCT 894
DB 3481 GAATGCAAGAACACAGAGAGGTTCTTACCAATGCCCTGTGACACAGGGGCTTCCAGGCTGTC 3540
QY 895 TATCTGAGATGATGATTAACCCGCTGTATGTCTCTGCTGAGAACCCCTGG 944
DB 3541 AATGGCACCATGTGTGAGAGCGTGAATGATGTGTGGGGAAGACCATTTG 3590

RESULT 15

PCT-US95-02251-2
Sequence 2, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patentln Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3753 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,722B
FILING DATE: 07-JUN-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US PCT/US95/02251
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100.000500/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7000
TELEFAX: (713) 934-7011
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3759
US-08-479-722B-3

```

```

Query Match      4.2%; Score 53.2; DB 3; Length 3759;
Best Local Similarity 49.4%; Pred. No. 2.1e-06;
Matches 167; Conservative 0; Mismatches 168; Indels 3; Gaps 1;

QY 466 TGTGGAGATGTCCTGAGATCCATTTCTGTACATGCAACCCGTTTACCCCTCAATGAG 525
DB 2215 TGTGAGAACCTTCGGGCTTCTTACCGTTCACGTGTGCCAGGAGATACGAACCCGACA 2274
QY 536 GATGGAAGGCTTGGCCAGATGTGAAGAGTGTGCCAGCAACCCCTGC---GTGCAA 582
DB 2275 GGACGGCTCAGTTGCATACGATGATGATGTGAGGCTGGAAAGTGTCCCAAGATGCC 2334
QY 583 ACCTGCGTCAACACCTACGCGCTCTTTCATCTGCGCTGTGACCCAGATATGAATTTGAG 642
DB 2335 ATCTGCACGACACACACGAGGCTCTTTCAGATGTCAAGTGTCTCCGGCTATCATCTGCA 2394
QY 643 GAAAGATGGGTTTCATGTGATGATGATGAGAGTGTGACAGTTCCTGTGAGTTCTTGCAA 702
DB 2395 AGGAGTCGAGCGCGCTGTGAGACATTTGATGATGATGATGATGATGATGATGATGATG 2454
QY 703 CATGATGTGTAAACAGCCCGGACATATCTTCTGCTCGCCCTCGAGCTCATATCTG 762
DB 2455 GGTGACTGATCATATACCAATGTTCTTCTACAGATGTCTGTGCTCCCTGGGTCATCGTTG 2514
QY 763 CTGATGACACCGAAGCTGCCAAGATCAACGAATG 800
DB 2515 GTGGGCGGACGAGGATGCAAGAAAGATATGATGAGTGTG 2552

```

```

RESULT 13
US-08-479-722B-1
; Sequence 1, Application us/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF $\beta$  BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Williams, Morgan & Amerson
STREET: 7676 Hillmont, Suite 250
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,722B
FILING DATE: 07-JUN-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US PCT/US95/02251
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100.000500/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7000
TELEFAX: (713) 934-7011
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5499 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5499
US-08-479-722B-1

```

```

Query Match      4.1%; Score 52.4; DB 3; Length 5499;
Best Local Similarity 47.4%; Pred. No. 4.8e-06;
Matches 223; Conservative 0; Mismatches 241; Indels 6; Gaps 2;

QY 481 GGATCCTATTTCTGTACATGCAACCCGTTTACCCCAATGAGATGGAAGTCTTGC 540
DB 3121 GGCTCTTATAGATGCTCTGTGAGCGGGCTATGAGTGTACCCCAAGAAAGGCTTC 3180
QY 541 CAGATGTGAAGAGTGTGCCAGCAGAGACCCCTG---CGTCAAACTGCGTCAACACC 597
DB 3181 CGAGATGTGAAGAGTGTGCCAGCAGAGACCCCTGCTGCTCCCAAGGCGCTGCTCAACAG 3240
QY 598 TAGGCTCTTTCATCTGCGG---CTGTGACCCAGATATGAATTTGAGAGAGATGCGCTT 654
DB 3241 GAGGGCTCTTTCACCTGCTGACGCTGTGAGAGCGGGTACTGGGTGAAGAAATGCACT 3300
QY 655 CATGCACTGATATGAGAGAGTGTGAGAGTCTTCTGTGAGTCTCTGCAACATGAGTGTG 714
DB 3301 GCCTGTGAAGACTTGATGAAATGTGCTTCTGAGTGTGCCCAAGAGGCTGTGCACC 3360
QY 715 AACGAGCCCGACATATCTTGTCTGCTGCGCTCCAGGCTACATCTCTGATGATCAAC 774
DB 3361 AATACTGTAGGCTCTCTTCTGCTGCAAGAGATGTACACAGGCGTACCGCCCAACCCCTG 3420
QY 775 CGAAGCTGCCAAGCATCAACGAATGTGAGCAGACAGAAACACACGTCGAACCTGAGCAG 834
DB 3421 GCCAAGAGATCGCAAGATGTGATGATGATGATGATGATGATGATGATGATGATGATG 834

```


Db	213	GCCATCAGACATCTTCAGATATACAGGCCAACATATTATTATGCACAAACCATCAATCAATCTT	272
QY	1068	CCAGATCAAAATCTTGGAAATGAGGCCAGAGATTTTACATGCGGCACAAAGGCCCCCATCAG	1127
Db	273	TCGGCTTAAATCTCTGGAAATGAAATGAG-AGTCTACTTACGAGAACCAAAACCCCTGTAA	331
QY	1128	TGCGACCCCTGTGATGACAGGCCCATCAAAAGGGCCCGGGAATCCAGCTGGACTTGG	1187
Db	332	TGCAATCTCTTGCTGCTGTAAGNCATATTATCAGGACCAAGAGACATATCGTGGACCTGGA	391
QY	1188	AATGATCAC	1196
Db	392	GATGCTGAC	400

```

RESULT 10
US-09-404-879A-340/C
: Sequence 340, Application US/09404879A
: Patent No. 6468546
: GENERAL INFORMATION:
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: King, Gordon E.
: APPLICANT: Algate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
: FILE REFERENCE: 210121.46202
: CURRENT APPLICATION NUMBER: US/09/404,879A
: CURRENT FILING DATE: 1999-09-24
: NUMBER OF SEQ ID NOS: 393
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 340
: LENGTH: 220
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(220)
: OTHER INFORMATION: n = A,T,C or G
: US-09-404-879A-340

```

Query Match	6.3%	Score 80.2	DB 4	Length 220
Best Local Similarity	60.5%	Pred. No. 1,4e-15		
Matches 130:	Conservative	0	Mismatches 85:	Indels 0
			Gaps 0:	
QY	298	TGTCGTGATGTGACGAGTGTGCAACAGATTCCACACAGTGCACACCCACAGATCTGC	357	
Db	219	TGTCGTGATGTGACGAGTGTGTCACAGGCCCTTGACGACTGTGCGCCACGACGAGTATGC	160	
QY	358	ATCAATACTGAAGCGGGGTACACCTGTCTCTTCGACACCGAGATATGTGGCTTCTGGAAGC	417	
Db	159	CATTAACCTGCCGTGCTCCTATCACTGACACCTGCCCTGTATGTTACCGGAAGATGGGCC	100	
QY	418	CAGTGTAGACATGATGATGTGCGTATGATGGTATCTCCACGACGCTCTTGCGAATGTT	477	
Db	99	GAGTGTGACATAGACGAGTGTGCGGTACCGGTACTACTCCAGACACCGCTGGGTGAACTG	40	
QY	478	CTTGATCTCTATTCTTGTACATGCAACCCCTGTT	512	
Db	39	CTGCGTCTCTCGCTCCGACAGNCCNAGCGGGCT	5	

RESULT 11
 US-08-282-141-1
 : Sequence 1, Application US/08028241
 : Patent No. 5538661
 : GENERAL INFORMATION:
 : APPLICANT: Schneider, Claudio
 : APPLICANT: Varnum, Brian
 : APPLICANT: Avanzzi, Giancarlo
 : APPLICANT: Brancolini, Claudio
 : APPLICANT: Manioletti, GuidoAlberto
 : TITLE OF INVENTION: Stimulating Factor for the AXL Receptor
 : NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESS: Amgen Inc.
STREET: 1640 Dehavenland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: United States
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/282,141
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2461 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
US-08-282-141-1

```

	Query Match	4.6%	Score 58.2	DB 1	Length 2461
	Best Local Similarity	54.5%	Pred. No. 4.5e-08		
	Matches 164	Conservative	0	Mismatches 120	Indels 9
					Gaps 2
QY	446 ATGGTACCTGCCACAGCTCTGTGCGAATGTTCTCGATCTTATTTCTGTACATGCAAC	505			
Db	625 ACGGGGGGTGCTCTCCAGATCTGCCACAAACAAAGCGGGGTAGCTTCCACTGTTCTCCACCA	684			
QY	506 CTGGTTTACCTCTAATGAGGATGGAGGTTTGGCAAGATGTGACGAGTGTGCCACCG	565			
Db	685 GCGGCTTCGACCTCTCCCTCTATGTGGCAGAGCTCTGCCAAGATATACACAGTGGCGAGACT	744			
QY	566 AGAACCCCTGGGTGCAAACCTGCCGTCACACCTA---CGGCTCTTTCATCTGGCCGCTGTG	622			
Db	745 CGGAGGCTTGGGGGAGGCGCGCTGCAAGAACCTGCCGGCTCTACTTCTGCTGTG	804			
QY	623 ACCCAGGATATGAATCTTGAGGAAAGATGGCTTCATTTGACGATATGGACGAGTGCACCT	682			
Db	805 ACGAGGGCTTTGGCTTACAGCTCCACGAGGAAGGCTTGGCGAGATGTGACCGAGT-----	859			
QY	683 TCTCTGAGTTCCTTGCCACATGATGTGTGAACACGCCCGGACATACTTCTGCTCT	742			
Db	860 -TCTCAGAGGGCGGCTGTGAGCAGGCTCTCGTGAACCTCCACGAGGAGTACACCTGCCACT	918			
QY	743 G 743				
Db	919 G 919				

RESULT 12
 US-08-479-722B-3
 Sequence 3, Application US/08479722B
 Patent No. 6074840
 GENERAL INFORMATION:
 APPLICANT: Bonadio, Jeffrey
 APPLICANT: Yin, Wushan
 TITLE OF INVENTION: LATENT TGF (BINDING PROTEIN (LTBP)
 TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Williams, Morgan & Amerson
 STREET: 7676 Hillmont, Suite 250
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77040
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

QY	Db
948	93
153	152
1008	1067

SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 259
 LENGTH: 1018
 TYPE: DNA
 ORGANISM: Human
 US-09-188-930-259

Query Match
 Best Local Similarity 62.3%; Pred. No. 2.4e-89;
 Matches 515; Conservative 0; Mismatches 309; Indels 3; Gaps 1;

24.9%; Score 316.6; DB 3; Length 1018;
 443 GCTATGTTACTCCAGCAGCTCTGTGCAATGTTCTGATTCCTATTCTGTATGCA 502
 1 GCTACCGCTACTGCAACAGCAGCTGCTGGAACCTGCTGCTGCTGCTGCTGCTGCTG 60
 503 ACCCTGTTTACCCTCAATGAGATGGAAGCTTGTGCAAGATGTGAACAGTGTGCA 562
 61 AGCCGGGCTTCCAGCTGGGCTTAAACAGCGCTCTGTTGATGTGAACGAGTGTGCA 120
 563 CCGAGAACCCCTGCTCAACCTGCTCAACAGCTGCTTCAATCTGCTGCTGCTGCTG 622
 121 TGGGGGCCCCATGAGCAGCAGCTGCTTCAATCTCTATGGAACCTTCTGCTGCTG 180
 623 ACCCAGATATGAACTTGAAGAGATGAGCTTCAATGCAATGATGAGCAGTCACT 682
 181 ACCAGGCTATGAGTCAATGAGATGAGCTTCTCTGCAATGATGATGATGATGAT 240
 683 TCTCTAGTCTCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGAT 742
 241 ACTCTAGTCTCTGCTGCAATGATGATGATGATGATGATGATGATGATGATGAT 300
 743 GCCCTCAGGCTCAATCTGCTGATGATGATGATGATGATGATGATGATGATGATG 802
 301 GCCCAGAGGTTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
 803 AGCAGAGAACACACAGCTGCAACCTGCAAGCAGAGCTGCTCAATTTTACAAGGGCTTCA 862
 358 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
 863 AATGATCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 922
 418 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477
 923 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 982
 478 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
 983 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
 538 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
 1043 GCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1102
 598 TCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
 1103 ACATGCGGCAACAGGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1162
 658 ACATGCGGCAACAGGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
 1163 CCCGGGAATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1222
 718 CCCGGGAATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
 1223 GCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
 778 CCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 824

RESULT 7
 US-09-248-757-1
 Sequence 1, Application us/09248757
 Patent No. 6417342
 GENERAL INFORMATION:

APPLICANT: STONE, EDWIN M.
 APPLICANT: SHEPHERD, VAL C.
 TITLE OF INVENTION: MACULAR DEGENERATION DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: UIA-018.02
 CURRENT APPLICATION NUMBER: US/09/248.757
 CURRENT FILING DATE: 1999-02-12
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 1
 LENGTH: 2512
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-248-757-1

Query Match
 Best Local Similarity 19.0%; Score 241; DB 4; Length 2512;
 Matches 547; Conservative 0; Mismatches 445; Indels 9; Gaps 3;

271 GATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 330
 403 GGTACAGAGCAAGTGAACAGCAACAGTGTGCAAGATGATGATGATGATGATGATG 462
 331 CACCAAGTGAACCCCAAGATGATGATGATGATGATGATGATGATGATGATGATG 390
 463 CACCAAGTGAACCCCAAGATGATGATGATGATGATGATGATGATGATGATGATG 522
 391 ACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
 523 CTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
 448 GGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507
 583 CCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
 508 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 567
 643 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
 568 AACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627
 703 AATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 762
 628 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 687
 763 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 822
 688 GATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 747
 823 AGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 882
 748 CCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807
 883 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 939
 808 AGGAGACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 867
 940 ACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 996
 868 ATGAGACCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
 997 TATCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1056
 928 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
 1057 CCAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
 988 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1047
 1117 ATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1176
 1048 CTTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1107
 1177 GCCAACCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1236

Oy 1162 CCCCCGGAAATCCAGCTGGAGCTTGGAAATGATCATCTGCAACAGCTGTCATCACTTCAGA 1221
 Db 1346 CCCCCGGAGTACGCTCTGAGCACTGGAGATGGTGCACATCAATGAAATTCCTCATATAGCTACCGG 1405
 Oy 1222 GGCAGCTCCGTGATCCGACTGGGATATATGTGTGCGAGTACCAATTC 1269
 Db 1406 GCGACGCTGTGACTGAGGCTACCGCTCTTGTATAGGGGGCTTCACACCTTC 1453

RESULT 5
US-08-980-514-2

; Sequence 2, Application US/08980514
; Patent No. 6004753

; GENERAL INFORMATION:

APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.

APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN S1-5-ECMP-LIKE PROTEIN

```

; TITLE OF INVENTION: EIL
;
; NUMBER OF SEQUENCES: 3

```

```
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

STREET: 3174 Porter Dr.
CITY: Palo Alto

STATE: CA
COUNTRY: USA

ZIP: 94304
COMPUTER READAB

MEDIUM TYPE: Diskette
COMPUTER: IBM compatible

```

; CONTROL: IBM compatible
; OPERATING SYSTEM: DOS
SOFTWARE: FASTSO for Windows

```

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: IIS/08/980 514

```

APPLICATION NUMBER: 05/08/9
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATE:
APPLICATION NUMBER:
ESTIMATED DATE:

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0436 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

```

; INFORMATION FOR SEQ ID NO:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 2018 base pairs
TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT13
;

```

CLONE: 2786449
US-08-980-514-2

Query Match

Best Local
Matches 61

malices of 262

ay	262
dy	530

Db	530
On	530

322 QY

Db 590

QY 382

Db 650

OY	442	CGCATGATGTTACTGCCAGACAGCTCTGTGCGAATGTTCCGTGATCTTATCTTTGATCAATGC	501
Db	710	CGCTAACGCGCTACTGCGACAGCACCGCGTGGCTGAACCTCGCTGGCTCTTCCGCTGCCAGATGC	769
OY	502	AACCCGTGTTTTACCCCTTCATATGAGAGTGAAGAGTCTTGTCCAAATGTGACAGATGTGCC	561
Db	770	GAGCGCGGGCTTCCAGACTGTGGGGCTCTTAACAAACCGCTCTCTGTGTGATGTGAACAGAGTGTAC	829
OY	562	ACCGAAGAACCCCTGCGTGGCCAAACGCTGGGTCAACACCTTACCGGCTCTTTCATCTGGCGGCTG	621
Db	830	ATGGGGGGCCCCATGCGAGACAGCGCTGCTTCAACTCTCTATGSGACCTTCTGTGTGCGCTGC	889
OY	622	GACCCAGATATGAACCTTGAGAGAAATGGCGCTTCATTGTCACTGATATGAGACAGATGACG	681
Db	890	CACCAAGGCGCTATGAGCTCATATCGGAGATGGCTTTCCTCTCATGTATATGATGAGTGTAGC	949
OY	682	TTCTCTAGTTCCTCTCTGCCAACAATGATGATGTGTGAACCAAGCCCGGCACATACTTCTGCTCC	741
Db	950	TACTTCACACTACTCTCTGTGCATATACCGCTGCGCTCAACGAGCCAGGCGCTTCTCTCCGAC	1009
OY	742	TGCCCTCCAGAGCTATACATCTGCTGGATGATGACAAACGGAACCTGCCAAGACATCAACGAATGT	801
Db	1010	TGCCACAGGGTTATACAGTCTGCTG---GCCACAGCGCTCTGCGCAAGACATATGATGATGTG	1066
OY	802	GAGCACAAGAACCAACAGCTGCACACCTGCAGACAGATGCTCAATTTACAAAGGGGGCTTC	861
Db	1067	GAGCTGTGTGGCGACACAGTGTCTCGAGGCGCCAAACCTGTGTCAACTTCCATCTCAAGGGGGCTTAC	1126
OY	862	AAATGCATCAACCCCATATCCGCTGTGAGAGGCGCTTATCTGAGATCAGTATACCGGCTGT	921
Db	1127	CGCTGCGCTGGACACCAACCGCTGCGTGGAGCCCTTACATCCAGTCTCTGAGAACCGGCTGT	1186
OY	922	ATGATGCTCTGTGAGAACCCCTGGGCTGCAGAGACCAAGCCCTTTACCATCTTGTACCGGGAC	981
Db	1187	CTGTGCCCCGGGCTCCAAACCTCTATATGTGAGAGAGCGCTTCATTCATATGTGCACCGGTAC	1246
OY	982	ATGACAGTGGTGTCAAGACGCTCCGCTTCCGCTGACATCTTCCAAATGCAAGCCACGAC	1041
Db	1247	ATGACACATCACCTCGGAGCGGAGGGGTGCCCTGACGTTGCCAGATCCAGCGACCTCC	1306
OY	1042	CGCTACCCCTGGGGCTATTATCAATTTCCAGATCAAAATGTGGGAATGAGAGGAGAAATTT	1101
Db	1307	GTTACACCCGGTGGCTCTCAATATGCTTTCAGATCCGCTGTGGAACCTGCGAGGGGAGCTTT	1366
OY	1102	TACATGGGGGCAAGGGGGCCCATCATGATGCCACCTGTGGATGACAGCGCCCATCAAAAGG	1161
Db	1367	TACATTTAGGCAAAATCAACAAAGTACAGCGCCATGTGTGGTCTCGCCCGGCGGGTACGGGC	1426
OY	1162	CCCCGGGGAATTCACAGCTGGAGCTTGGAAATGATCACTGTGCACACACTGTCATCAACTTGA	1221
Db	1427	CCCCGGGGGATACGTTGCTGAGACCTGAGATGTGTACCAATGAATTCCTCTATGAGTACCGG	1486
OY	1222	GGCAGCTCCGATCCGATCGACTGCGGATATATGTGTGCGAGTACCCATTC	1269
Db	1487	GCGAGCTCTGTACTGAGGCTCACCGTCTTGTATGAGGGGCTTACACCTTC	1534
RESULT 6			
US-09-188-930-259			
: Sequence 259, Application US/09188930A			
: Patent No. 6150502			
: GENERAL INFORMATION:			
: APPLICANT: Watson, James D.			
: APPLICANT: Strachan, Lorna			
: APPLICANT: Sleeman, Matthew			
: APPLICANT: Onrust, Rene			
: APPLICANT: Murlison, James Greg			
: TITLE OF INVENTION: Compositions and Methods Isolated from Skin Cells			
: TITLE OF INVENTION: and Methods for Their Use			
: FILE REFERENCE: 11000.1011c1			
: CURRENT APPLICATION NUMBER: US/09/188, 930A			
: NUMBER OF SEQ ID NOS: 348			
: CURRENT FILING DATE: 1998-11-09			

```

; CURRENT FILING DATE: 1998
; NUMBER OF SEQ ID NOS: 348

```

```

Db 487 GGCCTCAGGCTMCTCTTGTGAGATTAACCGAAGTGCAGAGATATCAATGATGTG 546
QY 803 AGCAGAGAAACACACAGCTGCACTGACGAGAGCTGTACAAATTTACAGGGGCTTGA 862
Db 547 AGCMCCGGAACACACATGACTCCCTGCGAGACTGTGTACATGTGACAGGGGCTTGA 606
QY 863 AATGATCGAACCCCATCCGCTGTGAGAGAGCTTATCTAGAGTACAGTGAATACCGCTGA 922
Db 607 AATGTATCGAACCCCATCCGCTGTGAGAGAGCTTATCTGTCTGTATTTGGGATACCGCTGA 666
QY 923 TGTGTCTGTGAGAACCTGTGCTGACAGACACAGCCCTTACATCTTGTACCGGACA 982
Db 667 TGTGCCCTGCTGAATACATGCTGTGACAGGAGACACCATTCATCTTGTTCGGACA 726
QY 983 TGGACGCTGTGAGAGAGCTGCCCTTCCCGCTGACATCTTCCAAATGACAGCAGACC 1042
Db 727 TGGATGTGTATCGAGAGCTGTGTCTGTCTGTGACATCTTCCAGATGCAAGGACGCC 786
QY 1043 GCTACCTGTGGGCTTATTAATTTCCAGATCAATCTGGGAATGAGGACAGAAATTT 1102
Db 787 GATACCTGGCGCTTATTAATTTCCAGATCAATCTGGGAGAGAGGCTGAGAGTTCT 846
QY 1103 ACATGCGGCAACGCGGCCCATCATGAGTGCACCTGTGTATGACAGAGCCCATCAAGGCG 1162
Db 847 ACATGCGGCAACGCGGCCCATCATGAGTGCACCTGTGTATGACAGAGCCCATCAAGGCG 906
QY 1163 CCGGGAATTCAGCTGAGTGTGAATGATCAGTGTCAACATCTCACTTCAGAG 1222
Db 907 CTGGGAGATTCAGCTGAGTGTGAATGATCAGTGTCAACATCTCACTTCAGAG 966
QY 1223 GCAGCTCCGTGATCGAGTGTGATATATGTGCGAGTACCATTC 1269
Db 967 GCAGCTCCGTGATCGAGTGTGATATATGTGCGAGTACCATTC 1013

RESULT 4
US-08-833-963C-1
; Sequence 1, Application US/08833963C
; Patent No. 5916769
; GENERAL INFORMATION:
; APPLICANT: Olsen, et al.
; TITLE OF INVENTION: Extracellular/Epidermal Growth Factor
; TITLE OF INVENTION: HCBAB58X
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Ave
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833, 963C
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US96/05033
; FILING DATE: 10-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 125..1453
US-08-833-963C-1

Query Match 29.3%; Score 371.2; DB 2; Length 1531;
Best Local Similarity 61.2%; Pred. No. 2.3e-106;
Matches 617; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

QY 262 TGGCGCTTGGATTCAGATGATGATGAAGCAACCAATGTGTGATGTGAGAGTGTGA 321
Db 449 TGGCCACAGGCTATGAGCCCGCATGAGACAGTGTGATGTGAGAGAGTGTGC 508
QY 322 ACAGATTCCACACAGTGCACACCCACAGATCTGATCAATACTGAAGCGGTACACC 381
Db 509 CAGGCCCTGACAGACTGTGCGCCACAGCAGACTGCCAATATCTGCTGCTATACAG 568
QY 382 TGTCTCTGACCGAGAGATTTGCTTGTGAGAGCCAGTCTTATGACATTTATGATAT 441
Db 569 TGCACCTGCTCCGTATGTGTACCGCAAGATCGGCGCCGAGTGTGTGACATGACGAGTGC 628
QY 442 CGCTATGTATACCGCAGACCTGTGTGCAATGTTCCTGTGATCTTCTTATCTGATAC 501
Db 629 CGCTACGCTATCTCCAGCAGCAGCTGTGTGAACTGCTGTGCTTCCGCTGACAGTGC 688
QY 502 AACCTGTGTTTACCTTCATGAGAGATGGAAGTCTTCCAGATGTGAAGAGTGTGC 561
Db 689 GAGCGGGGCTTCCAGCTGTGGGCTTAACAACGCTCTGTGTGTGATGTGAAGAGTGTGC 748
QY 562 ACCGAGAACCCCTGCTGCAAACTGCTCAACACTTACAGGCTTTATCTGCGCTGT 621
Db 749 ATGGGGGCGCCATCGAGCAGCAGCTGTCTCAACTCTATGAGACCTTCTGTGTGCTGC 808
QY 622 GACCAGAGATATGACTTGAAGAGATGAGGCTTATCTGATGATGATGAGAGAGTGCAGC 681
Db 809 CACGAGGCTATGAGTGTGATGAGGATGAGTGTCTCTGAGTATGATGATGATGATGATG 868
QY 682 TTTCTGAGTCTCTGTGCAATATGATGTGTGAACACCCGCGACATCTTCTGCTCC 741
Db 869 TACTCAGACTCTGTGCACTACCGCTGCGTCAACGAGCAGCGCGTCTCTGCTGCAAC 928
QY 742 TGCCCTCCAGCTTCAATCTCTGATGATGACAACCGAAGTGTGCAAGACATCAAGAAAT 801
Db 929 TGCCACAGGGTATACAGCTGCTG--GCCACACGCTCTGTGCAAGACATTAATGATGT 985
QY 802 GAGCAGAGAAACACACAGTGCACACCTGAGAGAGTGTCTATCAATTTACAGAGGCTTC 861
Db 986 GAGTCTGTGTGAGACACAGTGTGCGAGGCCCAACCTGTGTGATCACTTCATGAGGCTAC 1045
QY 862 AATGATCGAACCCCATCCGCTGTGAGAGAGCTTATCTGAGATGATGATTAACCGCTGT 921
Db 1046 CGTGGCTGTGACACCAACCGCTGTGAGAGCCCTATCAATCAAGTGTCTGAGAACCCGTGT 1105
QY 922 ATGTGTCTGTGTGAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
Db 1106 CTGTGCGCGCTCCCAACCTCTGTATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1165
QY 982 ATGAGAGTGTGTGAGAGAGCTGCTGCTGCTGATCTTCCAAATGAGAGAGAGAGAGAG 1041
Db 1166 ATGACATCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1225
QY 1042 CGCTACCTGGGCTTATTAATTTCCAGATCAATCTGGAATGAGAGAGAGAGATTT 1101
Db 1226 GTCTACCCCGGTGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1285
QY 1102 TACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1161
Db 1286 TACATTAAGCAATCAACAGCTGAGCGCATGTGTGTCTGCGCGCGGTGAGAGAGAG 1345

```

QY	61	TGCGGAACCATCCCGAGGCTGTGCGAGAGACATATGTGTATTACCAAAATGGCGGG	120
Db	517	TGCGGAACCATCCCGAGGCTGTGCGAGAGACATATGTGTATTACCAAAATGGCGGG	576
QY	121	TATTATGATATCCCGAGCAAAACCTGTGTATGCGAGGGCCCTACTCGAAGCCCTACTG	180
Db	577	TATTATGATATCCCGAGCAAAACCTGTGTATGCGAGGGCCCTACTCGAAGCCCTACTG	636
QY	181	ACCCCTACTCAGAGTCCGTACCCAGCAGCTGCCCCACACTCTCAGCTCCAACTATCC	240
Db	637	ACCCCTACTCAGAGTCCGTACCCAGCAGCTGCCCCACACTCTCAGCTCCAACTATCC	696
QY	241	ACGATCTCCAGGCGCTTTATATGCGCCTTGGATACCATGATGGATGAAGCAACATGT	300
Db	697	ACGATCTCCAGGCGCTTTATATGCGCCTTGGATACCATGATGGATGAAGCAACATGT	756
QY	301	GTGATGTGGAGAGTGTGCAACAGATTCCACAGTGCACACCCACAGATCTGTCAATC	360
Db	757	GTGATGTGGAGAGTGTGCAACAGATTCCACAGTGCACACCCACAGATCTGTCAATC	816
QY	361	AATACTGAAGGGGGGTACACCTGTCTCTGCAACGACGATATGTGCTTGTGAAGGCCAG	420
Db	817	AATACTGAAGGGGGGTACACCTGTCTCTGCAACGACGATATGTGCTTGTGAAGGCCAG	876
QY	421	TGCTTAGACATTTAGATATGTGCTATGGTATCTCCAGACAGCTGTGCGAATTTGCT	480
Db	877	TGCTTAGACATTTAGATATGTGCTATGGTATCTCCAGACAGCTGTGCGAATTTGCT	936
QY	481	GGATCCTATCTTGTACATGCACACCTGGTTTTTACCTCAATGAGAGTGAAGTCTTGC	540
Db	937	GGATCCTATCTTGTACATGCACACCTGGTTTTTACCTCAATGAGAGTGAAGTCTTGC	996
QY	541	CAGATGTGAAGAGTGTGCCACGAGAAACCCCTGCGCAACCTGGGTCAACACCTTAC	600
Db	997	CAGATGTGAAGAGTGTGCCACGAGAAACCCCTGCGCAACCTGGGTCAACACCTTAC	1056
QY	601	GGCTCTTTCATCTGCGCTGTGACCCAGATATGAATCTTGAGAGATGGGCTTCATTC	660
Db	1057	GGCTCTTTCATCTGCGCTGTGACCCAGATATGAATCTTGAGAGATGGGCTTCATTC	1116
QY	661	AGTATATGAGAGATGTGACGCTTCTCTAGTTCTCTGCGCAACATGACTGTGTGACAG	720
Db	1117	AGTATATGAGAGATGTGACGCTTCTCTAGTTCTCTGCGCAACATGACTGTGTGACAG	1176
QY	721	CCCGGACATATCTTGTGCTCTGCGCCTCAGAGCTCATCTGCTGAGTGAACCCGAGC	780
Db	1177	CCCGGACATATCTTGTGCTCTGCGCCTCAGAGCTCATCTGCTGAGTGAACCCGAGC	1236
QY	781	TGCCAAGACATCAGCAATGTGAGCACAGAAACACAGTCAACCTTGACAGACGTGC	840
Db	1237	TGCCAAGACATCAGCAATGTGAGCACAGAAACACAGTCAACCTTGACAGACGTGC	1296
QY	841	TACAATTTACAAGGGGGCTTCAAAATGCATGCACCCCATCCGCTGTGAGGAGACCTTATCTG	900
Db	1297	TACAATTTACAAGGGGGCTTCAAAATGCATGCACCCCATCCGCTGTGAGGAGACCTTATCTG	1356
QY	901	AGATCACTGATTAACCCGCTGTATGTCTCTGCTGAGAACCTTGCTGAGACACAGCC	960
Db	1357	AGATCACTGATTAACCCGCTGTATGTCTCTGCTGAGAACCTTGCTGAGACACAGCC	1416
QY	961	TTTACCATCTTGTACCGGGAGATGAGAGTGTGTACAGACGCTTCGTTCCCGTGTACATC	1020
Db	1417	TTTACCATCTTGTACCGGGAGATGAGAGTGTGTGTACAGACGCTTCGTTCCCGTGTACATC	1476
QY	1021	TTTCAAAATGCACAGCAGACCCGCTTACCCCTGGGCGCTATTACATTTTCCAGATCAAAATCT	1080
Db	1477	TTTCAAAATGCACAGCAGACCCGCTTACCCCTGGGCGCTATTACATTTTCCAGATCAAAATCT	1536
QY	1081	GGGAATGAGGCGAGAAATTTTACATCGCGCAACGGGGCCCATCAGTGTGCACCTGTGT	1140
Db	1537	GGGAATGAGGCGAGAAATTTTACATCGCGCAACGGGGCCCATCAGTGTGTGCACCTGTGT	1596

Accession	Sequence	Position
OY	ATGACAGCGCCCATCAAAAGGCGCCCGGAAATTCACGCTGGACTTGGAAATATATACATGTC	1200
OY		
Db	ATGACAGCGCCCATCAAAAGGCGCCCGGAAATTCACGCTGGACTTGGAAATATATACATGTC	1656
OY	1201 AACACTGTCATCAACTCAGAGGCGAGCTCCGTGATTCGACACGCGGATATATGTCTCCAG	1260
OY		
Db	1657 AACACTGTCATCAACTCAGAGGCGAGCTCCGTGATTCGACACGCGGATATATGTCTCCAG	1716
OY	1261 TACCCATTTC	1269
OY		
Db	1717 TACCCATTTC	1725

RESULT 3

US-09-188-930-67
; Sequence 67, Application US/09188930A

; GENERAL INFORMATION:

APPLICANT: Strachan, Lorna

```

; APPLICANT: Onrust, Rene
: APPLICANT: Murison James G

```

1111 TITLE OF INVENTION: Compositions Isolated From Skin Cells

CURRENT APPLICATION NUMBER: US/09/188,930A

NUMBER OF SEQ ID NOS: 348

! SEQ ID NO 67

TYPE: DNA

US-09-188-930-67

Query Match
Post Test

Matches	900;	Conservative	5;	Mismatches	102;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	------	--------	----	------	----

263 GCCGCTTTGGATACCGATGGATGAAGCAACCAATGTGTGGATGTGACGAGTGTCAA 322

5' GTCCTTGGGATCAGATGGAAGGCACCACTGTGTGTGGATGTGGACGAGTGTCCGA 66

220 CCGGATCCGATGCGATCCCGAGATCGCATCAATACGAGGCGGATCACCT 382

[illegible]

A vertical ruler with markings from 0 to 10 cm. The markings are in centimeters, with millimeter increments indicated between the centimeter numbers. The ruler is oriented vertically, with the 0 mark at the top and the 10 cm mark at the bottom.

2000

603 *****

247

OY 563' CCGAGGAACTCCCTGGCAACATCACCATTTCATAAACCTTCC

[illegible]

623 ACCCAGGATATGAACTTGAGGAAGATGCGCTTCATTGCAATGATATGGACGAATGCAGCTT 682

Db 367 ACCCAKGATATGAACTGGAGGAAGATGGCATTCACTGCABTGATATGGATGAGTGCABCT 426

683 TCTCTGAGTTCCTCTGCGCACATGAGTGTGTGAACGAGCCCGGCACATACTTCTGCTCCT 742

D22750.1
427 TCTCCGAGTTCCTCTGTCAACATGAGTGTGTGAACCAAGCCGGGCTCATACTTCTGCTCAT 486

743 GCCCTCAGGCTACATCCTGCTGGATGACAACCGAAGCTGCCAAGACATCAACGAATGTG 802

Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGATGA 60
Db 457 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGATGA 516
OY 61 TGGCGAACAATCCCGAGGCGTGGCGAGAGACATGATGTGTGTTAAACCAATGCGGG 120
Db 517 TGGCGAACAATCCCGAGGCGTGGCGAGAGACATGATGTGTGTTAAACCAATGCGGG 576
OY 121 TATTATGACATTTCCCGGACAAACCTCTGTATGAGAGGCGCTTACTGCAACCCCTACTG 180
Db 577 TATTATGACATTTCCCGGACAAACCTCTGTATGAGAGGCGCTTACTGCAACCCCTACTG 636
OY 181 ACCCCCTACTGAGGTCCGTACCCAGCAGTCCGCGCCACACTCTCAAGCTCAACTTCC 240
Db 637 ACCCCCTACTGAGGTCCGTACCCAGCAGTCCGCGCCACACTCTCAAGCTCAACTTCC 696
OY 241 ACATCTCCAGGCGCTTTATATGCGCTTTGATACAGATGATGAAAGCAACCAATGT 300
Db 697 ACATCTCCAGGCGCTTTATATGCGCTTTGATACAGATGATGAAAGCAACCAATGT 756
OY 301 GTGATGTGACAGTGTGCAACAGATTTCCACAGTGCACCCAGATCTGCATC 360
Db 757 GTGATGTGACAGTGTGCAACAGATTTCCACAGTGCACCCAGATCTGCATC 816
OY 361 AATACGAAAGGCGGTACACCTGCTCCGACCCAGGATATGCTTGGTTCGAAAGCCAG 420
Db 817 AATACGAAAGGCGGTACACCTGCTCCGACCCAGGATATGCTTGGTTCGAAAGCCAG 876
OY 421 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTCTGTGCGAATGTTCT 480
Db 877 TGCTTAGACATTTGATGATGTGCTATGCTACTGCGACAGCTCTGTGCGAATGTTCT 936
OY 481 GGATCTTATCTTTTACATGCAACCTCTGTTTACCTTCATGAGATGAAAGTCTTGC 540
Db 937 GGATCTTATCTTTTACATGCAACCTCTGTTTACCTTCATGAGATGAAAGTCTTGC 996
OY 541 CAAGATGTGAAGAGTGTGCGACGAGAAACCCCTGCTGCAAACTGCTCAACACTTAC 600
Db 997 CAAGATGTGAAGAGTGTGCGACGAGAAACCCCTGCTGCAAACTGCTCAACACTTAC 1056
OY 601 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGTGGCTTATTC 660
Db 1057 GGCTCTTTCATCTGCGCTGTGACCCAGATATGAACTTGAGAAAGTGGCTTATTC 1116
OY 661 AGTGAATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 1117 AGTGAATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
OY 721 CCCGCGACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1177 CCCGCGACATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1236
OY 781 TGGCAGACATCAAGAAATGTGAGCAGAAACACAGCTGCAACCTGAGAGAGCTGC 840
Db 1237 TGGCAGACATCAAGAAATGTGAGCAGAAACACAGCTGCAACCTGAGAGAGCTGC 1296
OY 841 TACAATTTACAAAGGCGCTTCAATGATGCAACCCCAATCCGCTGTGAGAGGCTTATTC 900
Db 1297 TACAATTTACAAAGGCGCTTCAATGATGCAACCCCAATCCGCTGTGAGAGGCTTATTC 1356
OY 901 AGATCAGTGTATACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 1357 AGATCAGTGTATACCGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
OY 961 TTATACATCTTTGACCGGACATGAGAGCTGTGTGAGAGCTGCTGCTGCTGCTGCTGCT 1020
Db 1417 TTATACATCTTTGACCGGACATGAGAGCTGTGTGAGAGCTGCTGCTGCTGCTGCTGCT 1476
OY 1021 TTCCAAATGCAAGCAGACCGCTGATACCTGGGCGCTTATTAATTTCCAGATCAATCT 1080
Db 1477 TTCCAAATGCAAGCAGACCGCTGATACCTGGGCGCTTATTAATTTCCAGATCAATCT 1536

```

```

OY 1081 GGAATAGGCGCAGAGAAATTTTACATGCGGCAAAACGGGCGCCATCAGTCCACCCCTG 1140
Db 1537 GGAATAGGCGCAGAGAAATTTTACATGCGGCAAAACGGGCGCCATCAGTCCACCCCTG 1596
OY 1141 ATGACAGCCCGCAAAAGGCGCGGGAATTCAGCTGAGCTTGAATGATCACTGTC 1200
Db 1597 ATGACAGCCCGCAAAAGGCGCGGGAATTCAGCTGAGCTTGAATGATCACTGTC 1656
OY 1201 AACACTGTCACTCACTTCAAGGCGAGCTCCGTGATCCGACTGCGGATATATGTGTCGAG 1260
Db 1657 AACACTGTCACTCACTTCAAGGCGAGCTCCGTGATCCGACTGCGGATATATGTGTCGAG 1716
OY 1261 TACCATTTC 1269
Db 1717 TACCATTTC 1725

```

RESULT 2
US-09-212-168-2
; Sequence-2: Application US/09212168
; Patent No. 6303765
; GENERAL- INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guebler, Karl J.
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/212,168
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/884,072
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PE-0333 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2550 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CORNMO701
; CLONE: 45517
; US-09-212-168-2

Query Match 100.0%; Score 1269; DB 4; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGATGA 60
Db 457 CAGTGCAGAAATGGCTTTGACCTGGATGCGCCAGTGCAGTCTTTTATGATTTGATGA 516

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 10:57:33 ; Search time 58.5574 Seconds

(without alignments)
6646.008 Million cell updates/sec

Title: US-09-674-379A-15

Sequence: 1 cagtcgcagatgcttga.....atgttcgcagtaaccattc 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

IssuedPatents: 1
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	2550	US-08-884-072-2	Sequence 2, Appli
2	1269	100.0	2550	US-09-212-168-2	Sequence 2, Appli
3	841.8	66.3	1260	US-09-188-930-67	Sequence 67, Appli
4	371.2	29.3	1531	US-08-833-963C-1	Sequence 1, Appli
5	371.2	29.3	2018	US-08-980-514-2	Sequence 2, Appli
6	316.6	24.9	1018	US-09-188-930-259	Sequence 259, App
7	241	19.0	2512	US-09-248-757-1	Sequence 1, Appli
8	106.8	8.4	3373	US-08-897-443-2	Sequence 2, Appli
9	89.4	7.0	401	US-09-643-597-278	Sequence 278, App
10	80.2	6.3	220	US-09-404-879A-340	Sequence 340, App
11	58.2	4.6	2461	US-08-282-141-1	Sequence 1, Appli
12	53.2	4.2	3759	US-08-479-722B-3	Sequence 3, Appli
13	52.4	4.1	5499	US-08-479-722B-1	Sequence 1, Appli
14	52.4	4.1	5502	PCT-US95-02251-17	Sequence 17, Appli
15	51.6	4.1	3753	PCT-US95-02251-2	Sequence 2, Appli
16	51.6	4.1	4314	US-08-199-780-2	Sequence 2, Appli
17	51.6	4.1	4314	US-08-316-650-2	Sequence 2, Appli
18	51.6	4.1	5089	5177197-31	Patent No. 5177197
19	50.4	4.0	6412	US-08-652-877-17	Sequence 17, Appli
20	50.4	4.0	6412	US-08-476-515A-17	Sequence 17, Appli
21	50.4	4.0	14042	US-08-652-877-85	Sequence 85, Appli
22	50.4	4.0	14042	US-08-652-877-89	Sequence 89, Appli
23	50.4	4.0	14080	US-08-652-877-87	Sequence 87, Appli
24	50.4	4.0	14083	US-08-476-515A-83	Sequence 83, Appli
25	45.6	4.0	14086	US-08-652-877-83	Sequence 83, Appli
26	45.6	3.6	3460	US-08-751-305-1	Sequence 1, Appli
27	44.8	3.5	2181	5208144-36	Patent No. 5208144

28	43.4	3.4	3546	US-08-872-757-3	Sequence 3, Appli
29	41.4	3.3	2452	US-09-381-779-1	Sequence 1, Appli
30	39.8	3.1	3156	US-09-284-819-8	Sequence 8, Appli
31	38	3.0	6048	US-09-634-920-3	Sequence 3, Appli
32	37.8	3.0	1611	US-09-249-697A-2	Sequence 2, Appli
33	37.8	3.0	1611	US-09-363-316B-2	Sequence 2, Appli
34	37.8	3.0	2365	US-09-249-697A-5	Sequence 5, Appli
35	37.8	3.0	2365	US-09-249-697A-18	Sequence 18, Appli
36	37.8	3.0	2365	US-09-363-316B-5	Sequence 5, Appli
37	37.8	3.0	2365	US-09-363-316B-23	Sequence 23, Appli
38	37.2	2.9	2397	5258288-2	Patent No. 5258288
39	37	2.9	6344	US-08-843-417-1	Sequence 1, Appli
40	35.4	2.8	825	US-08-312-870-6	Sequence 6, Appli
41	35.4	2.8	1338	US-08-307-444A-8	Sequence 8, Appli
42	35.4	2.8	1338	US-08-587-389-8	Sequence 8, Appli
43	35.4	2.8	1368	US-08-307-444A-7	Sequence 7, Appli
44	35.4	2.8	1368	US-08-587-389-7	Sequence 7, Appli
45	35.4	2.8	1425	US-08-307-444A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-884-072-2
Sequence 2 Application US/08884072
Patent No. 588234
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/884,072
FILING DATE: Herewith
CLASSIFICATION: A24
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0333 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2550 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CORNOTO1
CLONE: 45517
US-08-884-072-2
Query Match 100.0%; Score 1269; DB 2; Length 2550;
Best Local Similarity 100.0%; Pred. No. 0;